

Copyright

by

Adrian Garcia-Sierra

2007

**The Dissertation Committee for Adrian Garcia-Sierra certifies that this is the  
approved version of the following dissertation:**

**LANGUAGE CONTEXTS IN SPEECH CATEGORIZATION:  
TESTING THE DOUBLE PHONETIC STANDARD IN BILINGUALS**

**Committee:**

---

Craig A. Champlin, Supervisor

---

Harvey M. Sussman

---

Dennis McFadden

---

Randy L. Diehl

---

Swathi Kiran

---

Don C. Teas

**LANGUAGE CONTEXTS IN SPEECH CATEGORIZATION:  
TESTING THE DOUBLE PHONETIC STANDARD IN BILINGUALS**

**by**

**Adrian Garcia-Sierra, B.A.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**August 2007**

## **Dedication**

Dedicated to all graduate students that had made up their mind to seek simple answers  
from complicated questions

## **Acknowledgements**

I would like to thank my wife Nairán who supported me unconditionally throughout the achievements and pitfalls of grad school.

Thank you to my parents for always believing in my decisions and to my brother for laughing at my good and bad jokes.

Thank you to the committee dissertation members and to Edward Pasanen for acting as a guide in my research interests, and for providing me with feedback and support in designing experiments.

Thanks to all the research assistants, especially Jennifer Siard, who in some way or another made possible the present investigation.

Thanks to the McFadden lab for standing my mood changes while trying to survive grad school. The present document is proof that I survived.

**Language Contexts in Speech Categorization:  
Testing the Double Phonetic Standard in Bilinguals**

Publication No. \_\_\_\_\_

Adrián García Sierra, Ph.D.

The University of Texas at Austin, 2007

Supervisor: Craig A. Champlin

Speech sounds are typically perceived categorically. The acoustic information in speech sounds is perceptually grouped into phonetic categories. It is widely known that language influences the way speech sounds are categorized. That is, one's native language influences where category boundaries are placed. However, it is less understood how bilingual listeners categorize speech sounds. There is evidence showing that bilinguals have different category boundaries from monolinguals, but there is also evidence suggesting that bilinguals have different category boundaries depending on the language they are using at the moment. This phenomenon has been referred as the double phonetic boundary. The goal of this investigation was to verify the existence of the double phonemic boundary in bilingual listeners. As has been done in other studies, bilingual speakers of Spanish and English were asked to identify the speech sound /ta/

from a 10-token speech continuum ranging in VOT from /da/ to /ta/ in two language contexts. In this study, however, two additional procedures were carried out. First, English monolinguals were asked to identify the continuum in two language contexts. It was expected that bilinguals, but not monolinguals, would show a double phonetic boundary. Second, while participants' behavioral measures were assessed, electrophysiological measures [event-related potentials, (ERPs)] also were recorded. This was done in order to observe how speech sounds are represented in the brain. It was expected that bilinguals, but not monolinguals, would show different ERP amplitudes across language contexts. The behavioral results showed that phonemic boundaries did not differ across language contexts for either bilinguals or monolinguals. Further analyses showed bilinguals, but not monolinguals, perceived specific speech sounds—in the “ambiguous zone”—differently across language contexts. The electrophysiological results showed that the ERPs of bilinguals, but not monolinguals, differed across language contexts. Interestingly, behavioral measures correlated significantly with electrophysiological measures only in bilinguals. This result showed that the ERP amplitude was in accordance with the number of sounds perceived as ‘ta’ across language contexts. The challenges of testing the double phonemic boundary are discussed, along with the limitations of the methodology used in this study.

## Table of Contents

List of Tables .....	x
List of Figures .....	xi
Chapter 1 Introduction .....	1
Speech Categorization .....	1
Chapter 2 Literature Review .....	4
Behavioral Studies Assessing the Double phonemic boundary in Bilinguals	4
Electrophysiological Studies Assessing the Double phonemic boundary in Bilinguals .....	14
Summary: Literature Review .....	27
Chapter 3 Study Overview .....	29
Research Questions .....	33
Chapter 4 Method .....	36
Participants .....	36
Questionnaires .....	38
General Procedure .....	39
Behavioral and Electrophysiological Procedure .....	41
Chapter 5 Results .....	44
Behavioral Responses in Bilinguals and Monolinguals .....	44
Electrophysiological Analyses in Bilinguals and Monolinguals .....	51
Correlations between Electrophysiology and Behavior .....	60
Chapter 6 General Discussion .....	62
Speech Categorization and Bilinguals' Double Phonetic Boundary .....	62
Electrophysiology and Bilinguals' Double Phonetic Boundary .....	65
Correlation between Behavior and Electrophysiology .....	71
Limitations and Future Directions .....	72
Conclusion .....	74



Appendix A Language Questionnaire.....	76
Appendix B Stimulus Sequence.....	84
References.....	85
Vita .....	91

## **List of Tables**

Table 1:	Phonetic Representations of [t] and [d] Expressed in Duration of VOT .....	2
Table 2:	Bilinguals' and Monolinguals' Cumulative $d'$ Means in Two Language Contexts .....	50
Table 3:	Means and Standard Deviations for Standard and Deviant P2 and N2 ERP Components in Two Language Contexts.....	56

## List of Figures

Figure 1:	Percent difference in judging /ga/ between the first stimulus after precursor sentence and following the fifth stimulus after precursor sentence.....	13
Figure 2:	Hypothetical example of cumulative $d'$ values in two language contexts .....	30
Figure 3:	VOT values from which phonetic decisions and ERP responses were obtained in this study .....	31
Figure 4:	ERPs from a bilingual in both language contexts.....	34
Figure 5:	Bilinguals' and monolinguals' phonemic boundaries (50% point) in two language contexts derived from curves fitted via PROBIT analysis .....	44
Figure 6:	Bilinguals' and monolinguals' percent of hit responses in two language contexts as a function of VOT .....	46
Figure 7:	Bilinguals' and monolinguals' cumulative $d'$ in two language contexts as a function of stimulus number .....	48
Figure 8:	Bilinguals' ERP response associated with standard and deviant sounds in two language contexts.....	52
Figure 9:	Bilinguals' and monolinguals' P2 and N2 deviant ERP components in both language contexts.....	56
Figure 10:	Bilinguals' and monolinguals' MMN amplitudes in both language contexts. ....	58
Figure 11:	Behavior and eERP scatterplots.....	60

## **Chapter 1: Introduction**

### **SPEECH CATEGORIZATION**

Speech sounds are typically perceived categorically (Liberman, Harris, Kinney, & Lane, 1961). The acoustic information in speech sounds is perceptually grouped into phonemic categories. Most of the time an acoustic feature is perceptually exploited to group speech sounds into categories. For example, Abramson and Lisker (1967) demonstrated that the interval between the release of the articulatory occlusion and the onset of voicing (voice onset time or VOT) differentiates /b-p/, /d-t/, and /g-k/. Also, they showed that speakers of different languages categorize speech sounds differently. Accordingly, Abramson and Lisker (1967) synthesized a total of 37 speech sounds, varying in physically equal VOT steps, and presented the speech sounds to monolingual speakers of different languages. The results showed that listeners grouped the speech sounds into no more than three phoneme categories. Interestingly, the listeners' native language influenced where the category boundaries were placed. These findings suggested two things: First, the ability to discriminate a set of stimuli varying along one dimension is limited to the ability to identify them as different sounds (Liberman et al., 1961). Second, one's native language influences the way speech sounds are categorized. Indeed, other studies support the assumption that one's native language deeply influences the way speech sounds are perceived (Best, McRoberts, & Sithole, 1988; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992).

The fact that a person's native language can influence speech perception suggests that across languages, identical phonetic information does not necessarily represent the same phonemic sound (Abramson & Lisker, 1967, 1972; Lisker & Abramson, 1964). For

example, Spanish phonemes /b/, /d/ and, /g/ (voiced stop consonants) are produced with a long lead (or prevoicing), as shown by negative VOT values of -100 ms or more, and voiceless stop consonants (Spanish phonemes /p/, /t/ and, /k/) are produced with short-lag (0 to +25 ms of VOT). By comparison, English phonemes /b/, /d/, and /g/ (voiced stops) are produced with positive VOT (0 to +25 ms of VOT), and voiceless stops (English phonemes /p/, /t/ and, /k/) are produced with long lag (or aspiration) as indicated by positive VOT values from +30 up to +100 ms (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Williams, 1977) (see Table 1).

Table 1: Phonetic Representations of [t] and [d] Expressed as Duration of VOT

VOT (ms)	-100	-90	-80	-70	-60	-50	-40	-30	-20	-10	0	5	10	15	20	25	30	40	50	60	70	80	90	100
Phonetic	[d]										[t]													
Phonemic EN											/d/										/t/			
Phonemic SP	/d/										/t/													

Phonetic categories are marked with [ ] and phonemic English (EN) and Spanish (SP) categories are marked in / /. Gray indicates the VOT values that are not commonly present in English or Spanish.

Each language assigns the degree to which a given phonetic feature (i.e., duration of VOT) will be used to describe a phonemic category or class. Commonly, and in order to avoid confusion, phonetic categories are represented with square brackets [ ] and phonemic categories are represented with forward slashes / /. English uses the [t] phonetic category to represent /d/ and /t/ phonemic sounds, and Spanish uses the [t] phonetic category to represent the phonemic sounds /t/ (Keating, 1984). This means that the same duration of VOT represents English /d/ and Spanish /t/. The Spanish /d/, in the other hand, is represented with negative VOT (pre-voicing), and thus it belongs to a phonetic category that is not commonly used in English ([d]).

The goal of the present investigation is to better understand whether bilingual speakers of English and Spanish have developed a double phonemic boundary that helps

them disambiguate the phonemic English /d/ from the phonemic Spanish /t/. Behavioral evidence concerning the double phonemic boundary in bilinguals will be presented first, and electrophysiological evidence concerning speech categorization in two language contexts will be presented next.

## **Chapter 2: Literature Review**

### **BEHAVIORAL STUDIES ASSESSING THE DOUBLE PHONEMIC BOUNDARY IN BILINGUALS**

Spanish and English represent different phonemic categories by means of the same phonetic category. Spanish /t/ and English /d/ are represented by [t] phonetic category. Therefore, Spanish /t/ is more likely to be perceived as a /d/ by monolingual speakers of English. Bilingual speakers of these languages need to learn to disambiguate the fact that identical sounds represent more than one phonemic sound. Bilinguals seem to have little difficulty learning to allocate identical phonetic information to different phonemic categories. Bilingual speakers of Spanish and English might be able to assign identical sounds to more than one phonemic category depending on the context. If this is the case, bilinguals of these languages would rely on an English or Spanish phonemic boundary depending on the language context. This assumption has been tested and it is known as the double phonemic boundary in bilinguals (Elman, Diehl, & Buchwald, 1977; Flege & Eefting, 1987a; Hazan & Boulakia, 1993).

#### **Double Phonetic Boundary**

Caramazza et al. (1973) investigated the placement of phonemic boundaries by bilinguals when categorizing speech sounds in two language contexts. It was assumed that any indication of a double phonetic representation would be reflected in the position of the phonetic boundary. Specifically, they proposed that the perceptual boundary dividing a voiced-voiceless continuum should be in agreement with the language a bilingual was using at the moment. In that study, bilingual speakers of French and English were asked to identify the same set of speech sounds twice. In one experimental session, English was emphasized by having a brief conversation in English before the experiment, whereas in the second session the conversation occurred in French.

Monolinguals, on the other hand, only were exposed to English conversations before the identification task. The results showed no differences in bilinguals' perceptual boundaries across language contexts; however, bilinguals' phonemic boundaries were at intermediate VOT values from those of monolingual speakers. These findings suggested that bilinguals map acoustic information onto speech categories by means of a phonetic criterion that encompasses the phonetic rules of both languages. Interestingly, the results did not reveal evidence of a double phonemic boundary in bilinguals (for French-English bilinguals at least). The missing evidence for a double phonemic boundary in bilinguals could be the consequence of not having established a language context during the actual perceptual task. Participants were engaged in conversations and asked to produce words in English or French before the perceptual task, but during the perceptual task, neither English nor French was emphasized. Therefore, bilinguals could have shifted back and forth between French and English phonetic rules in an uncontrolled way when judging speech sounds. If this were the case, an average phonemic boundary with intermediate VOT values from those of monolingual speakers would be expected.

Similar results to Caramazza et al. (1973) were reported by Williams (1977) who obtained speech judgments from 8 bilingual speakers of Spanish and English. She investigated the effect of phonemic boundaries when categorizing speech sounds in Spanish or English contexts. Language contexts were established by 10-min conversations in the language of interest before the experimental task. The results showed that only two of the participants had moved their phonemic boundary closer to the monolingual value for that language. The rest of the participants showed no displacement of the perceptual boundary across language contexts. Therefore, it seems



that some bilinguals were unable to perceptually separate Spanish and English phonetic categories, probably because they might have merged both categories into a single one. The author mentioned that her findings should not be taken as proof of the absence of a double phonetic standard, rather they only indicated that the conditions of the experiment did not elicit such an effect. In fact, after the 10-min conversation, Spanish or English was not emphasized during the perceptual task.

These two studies addressed the issue of bilinguals' double phonetic standard through language contexts. The primary method to create language contexts was to interact with participants in the language of interest before the perceptual task. Conversations and/or word production in the relevant language were done before the task, but language contexts were not established during the identification task. Elman et al. (1977) suggested that bilinguals needed to stay focused on the language of interest throughout the entire experiment. Therefore, these researchers delivered precursor sentences in the language of interest during the actual perceptual task. In their experiment, an exclusive precursor sentence was delivered before the presentation of each stimulus (*Escriba la palabra \_\_\_\_* or *Write the word \_\_\_\_*). Elman et al. assumed that bilinguals would exploit precursor sentences' phonetic information to disambiguate language-phonetic similarities. The results showed that bilinguals did assign identical acoustic tokens to different phonemic categories, depending on the language they were hearing at the moment. Moreover, Elman et al.'s data showed that the perceptual switching increased as bilinguals' proficiency in the second language (L2) increased. These findings showed that 1) the use of precursor sentences during phonetic decisions

seemed to help bilinguals stay focused on the language of interest, and 2) that the shift in the phonemic boundary depends on the degree of bilingualism.

In a relevant study, Flege and Eefting (1987a) investigated whether the magnitude of the shift in the perceptual boundary reported in bilinguals depended on L2 proficiency. Dutch speakers of English were assessed in L2 proficiency by accent ratings, by self-reports of proficiency, and by acoustic analyses of productions in both languages.

Language contexts were set by presenting pre-recorded instructions of the perceptual task in Dutch or English. In contrast with the studies summarized above, Flege and Eefting did not have conversations with the participants, rather the interactions with participants were minimized as much as possible. The experiment consisted of identifying a speech continuum ranging in 16 equal VOT steps from /da/ to /ta/. In addition to the speech tokens, participants heard 27 pre-recorded questions during the behavioral task. A single block consisted of 10 speech tokens followed by a set of questions in the relevant language. The results did show a change in bilinguals' phonemic boundary in accordance with the language context, although the magnitude of the phonemic boundary shift between language contexts did not vary as a function of L2 proficiency level. In fact, bilinguals that scored low in the L2 proficiency scale and had strong foreign accent (i.e., word productions) showed the strongest voicing boundary shift. The same results were found in five native English speakers who could speak Dutch to some extent. Although, the findings were in contradiction with the Elman et al. (1977) study that found that perceptual switching increased with L2 proficiency, the production data were consistent with the Elman et al. study. That is, Dutch speakers of English showed VOT realizations in accordance with their L2 foreign accent scores. Specifically, bilinguals who spoke

English with a strong Dutch accent (low score in the foreign-accent-scale) showed small VOT values when producing English /t/ (see also Flege & Eefting, 1987b). That is to say the English /t/ was produced like the Dutch /t/.

In a more recent study, Hazan and Boulakia (1993) investigated whether bilinguals could weight speech cues in accordance to their native language (L1). That is, bilinguals might use speech cues that are relevant in L1 when using L1, but not when using L2. The authors synthesized two VOT continua ranging from /b/ to /p/. Two continuums were synthesized. In one continuum the speech stimuli finalized with the /en/ part of the word *pen*. In the other continuum, the stimuli finalized with the /en/ part of the word *Ben* (/ɛ/ vowel sound). In this way, both continua had allophonic sounds, but each continuum included different spectral energy at /e/ onset. That is, the *Ben* continuum had lower F1 spectral energy at /e/ onset than the *pen* continuum. Large negative VOT values and low F1 at /ɛ/ onset are not representative speech cues in English, but both cues are permissible in French. Hazan and Boulakia investigated the extent to which the identification functions varied across different language contexts as well as bilinguals' sensitivity to other cues that have a different perceptual weight in English and French.

In the Hazan and Boulakia (1993) study, bilinguals were tested twice in separate sessions. In each session, only one language was used. Each session started with recording bilinguals' productions of sentences and minimal pairs (e.g., Ben and pen or Benne and Peine) for later evaluation of foreign accent. Language contexts were set by engaging participants in a conversation and by filling out a questionnaire in the language of interest. During the actual perceptual task, all speech tokens were preceded by the

word “repeat” in English or “répète” in French. The words were used to establish the language context of interest throughout the experimental task. The results showed that French-dominant and English-dominant bilinguals had a shift in the phonemic boundary in accordance with language context for both VOT continua. Further, the results revealed little evidence that bilinguals acquiring L2 at earlier ages produced greater shifts in the phonetic boundary. However, there was a tendency for more early L2 learners to show the shift in the phonemic boundary than late L2 learners. Regarding VOT production measurements, bilinguals produced less +VOT in the French than in the English language context. Finally, in reference to bilinguals’ sensitivity to cues not present in their native language, it was shown that bilinguals were able to implement phonetic rules not allowed in their native language. For example, monolingual English speakers had difficulty judging /p/ sounds in the *Ben* continuum, but English-dominant bilinguals did not have the same difficulty.

### **Summary of Studies Testing the Double phonemic boundary in Bilinguals**

The experiments discussed thus far reveal different results depending on the method implemented to test bilinguals’ double phonetic standard. The three experiments that did show a shift in bilinguals’ double phonemic boundary presented acoustic information (i.e., PS, questions or words) in the language of interest during the behavioral task. However, those investigations did not agree in the way L2 proficiency predicted the amount of shift in the voicing boundary. In the Elman et al. (1977) investigation, bilinguals were interviewed individually and the experimenter rated bilinguals’ language fluency on a scale ranging from 0 to 8. In the case of Flege and Eefing’s (1987a) experiment, bilinguals’ L2 proficiency was assessed by accent ratings, self-reports of

proficiency, and acoustic analyses of productions in both languages. In Hazan and Boulakia's (1993) experiment, a questionnaire that assessed proficiency in both languages was used, and bilinguals were asked to read out loud "accent-revealing" sentences. In the Elman et al. study, the shift in the phonemic boundary was correlated with the bilinguals' L2 proficiency. Flege and Eefing showed exactly the opposite outcome. That is, the less proficient L2 speakers showed the strongest shift in the phonetic boundary. Hazan and Boulakia reported that early bilinguals showed only a tendency to have larger shifts in the phonetic boundary. The lack of agreement in these investigations may be the consequence of two factors. First, language contexts might not have been sufficiently strong to establish a specific language context in all participants. Second, the variability encountered in participants' L2 proficiency makes it difficult to establish what a true bilingual speaker is. Unfortunately, improving the way language contexts are established seems to be the only variable that can be objectively improved because finding bilinguals who are equally fluent in both languages is the exception rather than the rule (Grosjean, 1982)

### **Limitations of Bilingual Studies and Speech Categorization**

One of the most important limitations of the previous studies is that monolingual speakers were only tested in their native language. Indeed, there is only one investigation that has assessed monolinguals' phonetic decision in two language contexts (Bohn & Flege, 1993). Bohn and Flege (1993) investigated monolinguals' phonetic decisions in English and Spanish language contexts. The authors wondered if monolinguals would show a voicing boundary shift in accordance with the language being used in the acoustic context. Specifically, they assessed bilingual speakers of English and Spanish and

monolingual speakers of English in two language contexts. Both groups were asked to identify nine /t/ short lags (short VOT) produced by monolingual speakers of Spanish (Spanish /t/). The stimuli were presented with a 1-s interstimulus interval and the sounds were followed by the carrier phrase “\_\_ es la palabra” in the Spanish language context or by “\_\_ is the word” in the English language context. The results showed that language contexts did affect monolinguals in the same way as bilinguals. Both groups judged more short-lag speech sounds as /t/ sounds in the Spanish than in the English language context. Bohn and Flege concluded that the carrier phrases played the role of ‘acoustic adaptors’ rather than a context for the language under test. During the speech identification task it is possible that the acoustic information in the carrier phrases affected the perceptual judgment given to the speech stimuli. Indeed, Holt (2005) showed that the phonemic boundaries obtained from speech categorization tasks can be affected by the acoustic context preceding the phonetic decisions. She reported a significant voicing boundary shift even when the speech sounds and the acoustic context were separated by as much as 1300 ms.

Garcia-Sierra and Champlin (2003) assessed bilinguals’ phonemic boundary shift using a method that considered the effect of precursor sentences. Bilingual speakers of Spanish and English and monolingual speakers of English were asked to identify voiced and voiceless sounds from a continuum ranging from /ga/ to /ka/. The speech continuum consisted of 27 speech tokens that were delivered 10 times each. Language contexts were established by the experimenter interacting in Spanish or English with the participants and by presenting precursor sentences (in the language of interest) before the speech token to be identified. Precursor sentences were presented 13% of the times in a

random order before the presentation of the speech stimulus to be identified as /ga/ or /ka/. The remaining 87% of the stimulus presentations occurred without precursor sentences. It was expected that reducing the number of precursor sentences presentation would help avoiding biases in participants' phonetic judgments. Garcia-Sierra and Champlin calculated the phonemic boundary shift of each participant in both language contexts and the degree to which precursor sentences produced an effect in participant's judgments. The results did show a shift in the phonemic boundary in accordance with the language context in the bilingual group, but not in the monolingual group; however, both groups were affected by precursor sentences. Garcia-Sierra and Champlin determined the effect produced by precursor sentences by calculating the percent of times stimuli occurring in first to fifth position after filler sentences were judged as /ga/. The percent difference between first and second, first and third, first and fourth, and first and fifth positions was used as an estimate of change in judging /ga/.

Figure 1 shows the percent difference for the first five judgments occurring after precursor sentences for stimuli with 0, 5, 10, 15, 20, 25, and 35 ms of VOT. Stimuli in second to fifth positions were judged differently from stimuli in the first position. Specifically, a positive percent difference indicates that the first position was judged as /ga/ and positions from second to fifth were judged as /ka/. A negative percent difference indicates the opposite outcome; the first position was judged as /ka/ and positions from second to fifth were judged as /ga/. Values of zero indicate that the same stimulus was judged as /ga/ regardless if presented immediately after a precursor sentences or in following positions after a precursor sentence.

Figure 1 shows that judgments for both groups were biased by precursor sentences. Specifically, the percent difference between first and fourth and first and fifth positions in the monolingual group differed between language contexts. These results are in agreement with Bohn and Flege (1993) who reported bilinguals' shift in the voicing boundary was more likely the consequence of filler sentences used to establish the language contexts than bilinguals' showing a double phonetic boundary. It is interesting to note that in Garcia-Sierra and Champlin's (2003) experiment both groups' stimuli comparisons did not change over time. More precisely, only the judgments given to stimuli in first position were biased by the acoustic information in the precursor sentences. Research has shown that the effect acoustic precursors exert over phonetic decisions diminishes within several hundred milliseconds (Holt & Lotto, 2002; Lotto, Sullivan, & Holt, 2003; see Holt, 2005, for times exceeding 1 sec). Although it is clear that precursor sentences affected bilinguals' and monolinguals' decisions in the Garcia-Sierra and Champlin study, it is also evident that the effect was less than if a sentence had been presented before each phonetic judgment.

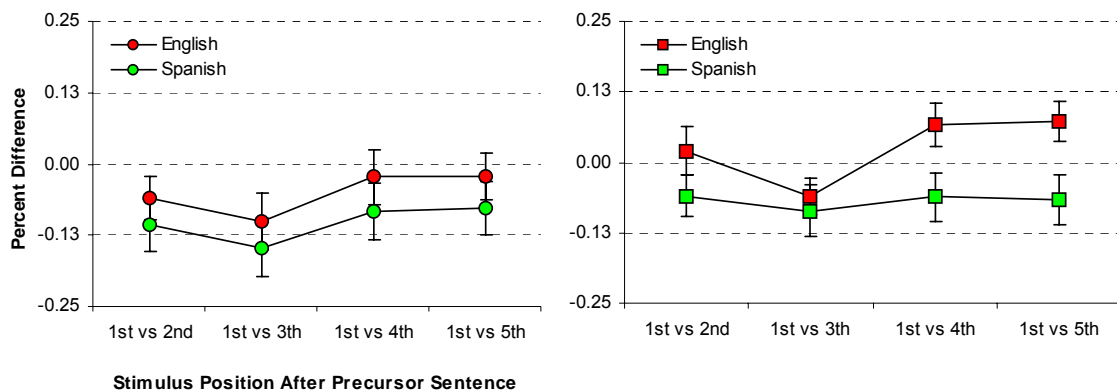


Figure 1: Percent difference in judging /ga/ between first stimulus after precursor sentence and fifth stimulus following precursor sentence. Note: Left (bilingual) and right (monolingual); panels show the percent difference in both language contexts.



Garcia-Sierra and Champlin's (2003) investigation gave evidence for bilinguals' double phonetic boundary. More research is needed to test other methods for establishing language contexts. It appears that the presentation of sentences in the language of interest during the perceptual task is not the best way to set language contexts.

## **ELECTROPHYSIOLOGICAL STUDIES ASSESSING THE DOUBLE PHONEMIC BOUNDARY IN BILINGUALS**

### **Auditory Evoked Potentials**

Human Auditory Evoked Potentials (AEPs) recorded from the scalp reflect transient neural changes that progress from the cochlea to the cortex. AEPs give detailed temporal information that can be used to establish objective measures in auditory neurological disorders (Picton, Hillyard, Krausz, & Galambos, 1974). AEPs are miniscule changes caused by sensory stimuli in the electroencephalogram (EEG) that are normally buried by different spontaneous brain rhythms. These physiological responses to stimulation can be seen by using various techniques to remove unwanted noise (Donchin, Ritter, & McCallum, 1978).

The AEPs presumably reflect populations of neurons firing synchronously in response to sounds, and represent the obligatory auditory-brain responses to auditory stimuli. Fifteen individual components have been described for the auditory modality by means of AEPs (Picton et al., 1974). The 15 components are divided into early-latency, middle-latency and long-latency responses based on their appearance following stimulation. The first seven (I to VII) early components occur between 1 to 12 ms after stimulus onset. There is good evidence to suggest that the early components are generated by the auditory nerve fibers and auditory brainstem structures through the lateral lemniscus and inferior colliculus (Altenmüller & Gerloff, 1998). The middle-

latency components ( $N_o$ ,  $P_o$ ,  $N_a$ ,  $P_a$ , and  $N_b$ ) occur at latencies from 12.5 to 50 ms after stimulus onset. The possible auditory structures giving rise to these components include the thalamus, primary auditory cortex, and scalp musculature. The last four evoked potentials ( $P_1$ ,  $N_1$ ,  $P_2$ , and  $N_2$ ) are called the late components or long-latency components. The long-latency components have a widespread distribution over the frontal scalp area, and they occur at latencies between 50 to 500 ms after stimulus onset. Research suggests that long-latency components most likely originate from the cortex (Picton & Hillyard, 1974).

The 15 individual auditory evoked potentials represent the volley of electric activity associated with acoustic stimulation. Considerable research has been done to better understand the evoked activity that is associated with sensory integration and the electric activity associated with pre-cognitive and cognitive processes. It is well accepted that exogenous AEPs (e.g., brainstem auditory evoked potentials) mainly depend on physical properties of the stimulus (Hall, 1992). For example, it has been shown that there is no significant change in the brainstem auditory potential amplitude when attention is directed toward the auditory stimulus (Picton & Hillyard, 1974).

### **Event Related Potentials**

Event Related Potentials (ERPs) represent changes in electric activity that occurs in response to sensory, cognitive, or motor events, and their response depends on context and instructions (Altenmüller & Gerloff, 1998). Investigators have attempted to match electrophysiological responses with psychophysical responses obtained in behavioral tasks with the purpose of better understanding the neural mechanisms underlying human attention and perception. Certain components within the ERPs are associated with types

of perceptual and psychophysical judgments such as detection, discrimination, and interpretation of sensory input (Osterhout, McLaughlin, & Bersick, 1997). The present investigation will use an ERP that has been used to study pre-attentive sound discrimination. The ERP is known as the mismatch negativity or MMN. A brief summary of MMN is given next.

### **Event Related Potentials Involving Sound Discrimination: The Mismatch Negativity**

The mismatch negativity (MMN) reflects changes in electric brain activity associated with pre-attentive sound discrimination or acoustic change detection (Naatanen, Gaillard, & Mantysalo, 1978; Naatanen & Michie, 1979). According to Naatanen (1982; 1992), the MMN represents the brain's automatic processes involved in encoding a stimulus difference or change. The MMN data suggest that a stimulus leaves a short-duration trace representing only physical stimulus features in the sensory-specific cortex (Naatanen, 1992). Sound discrimination is accomplished by comparing more than one neural representation or "memory trace."

The MMN is generally obtained when infrequent stimuli are randomly embedded in a train of frequently occurring stimuli. The MMN component is inversely related in amplitude to stimulus probability and is modality specific in its scalp distribution. It can be elicited whether the infrequent stimulus changes are attended to or ignored (Naatanen, Simpson, & Loveless, 1982). This means that the neural-mismatch between frequent and infrequent sounds is pre-perceptual and independent of attention and stimulus significance (Naatanen, 1986).

The odd-ball paradigm is used to acquire the MMN. The odd-ball paradigm consists of delivering two or more unique sounds in random order. One sound functions

as a template (standard) that is delivered 80-90% of the time and the other sound (target or deviant) is delivered 10-20% of the time. Neural activity is recorded from the standard and deviant sounds in two separate buffers that are averaged to yield two ERPs. The ERP associated with the standard stimulus conveys the sensory information associated with the standard sound, while the ERP associated with the deviant stimulus conveys the sensory information from the deviant sound plus neural activity associated with sound change detection. This “extra” neural activity is manifested by an increase of amplitude in the P<sub>2</sub>/N<sub>2</sub> complex in the ERP waveform. The amplitude differences in the standard and deviant P<sub>2</sub>/N<sub>2</sub> complexes are easier to see when subtracting the standard ERP waveform from the deviant ERP waveform. The resulting difference waveform is called the “mismatch negativity” and the most prominent negative peak occurs around 200 ms after stimulus onset. The amplitude of the MMN increases as the acoustic differences between standard and deviant increase (Tiitinen, May, Reinikainen, & Naatanen, 1994).

### ***Mechanisms Underlying Auditory Brain Discrimination***

Naatanen and Winkler (1999) have suggested that in order to acoustically differentiate one sound from another, the auditory system must fully resolve the acoustic features of a signal. Naatanen and Winkler proposed a series of properties that define the neural representation of sound. Indeed, they hypothesized that the neural activity undergoes a series of rapid transformations in which auditory features are extracted. For example, subcortical pitch-specific neurons respond as rapidly as 15 ms after stimulus onset (Greenberg, Marsh, Brown, & Smith, 1987) whereas loudness summation can be recorded approximately 150-200 ms after stimulus onset (Scharf & Houstsna, 1986). These findings and many others have shown that the acoustic features of sound are

integrated in different brain areas at different moments of time. Naatanen and Winkler suggested that the neural information being carried in the afferent activating pattern must be blended into a comprehensive neural representation in order to acoustically differentiate one sound from another.

Naatanen and Winkler (1999) proposed that sound representation should occur at the time when all the acoustic features extracted by the afferent activating pattern are integrated into a unitary percept. According to Naatanen and Winkler, memory plays an important role during the formation of sound representation in the brain. The auditory system creates individual memory traces of the acoustic features being processed at different moments in time. These memory traces are called *sensory feature traces* and are not yet part of the conscious experience. Once all the stimulus features exist in the form of *sensory feature traces*, the information can be merged into an *auditory stimulus representation* (see Balota & Ducheck, 1986; Cowan, 1995 for further discussion). Therefore, the emergence of a conscious percept is intimately linked to the formation of the sensory stimulus representation, and an *auditory stimulus representation* must exist in order for discrimination to occur.

It is important to mention here that the MMN does not reflect auditory stimulus representation, rather it reflects the mechanisms underlying sound discrimination or acoustic-change detection (Naatanen, Jacobsen, & Winkler, 2005; Naatanen, Pakarinen, Rinne, & Takegata, 2004). According to Naatanen and Winkler (1999), the MMN is an indirect measure of sound representation since an auditory stimulus representation must exist in order for discrimination to occur. Therefore, it has been said that the mechanisms

underlying the generation of the MMN reflect the brain's automatic change-detection response (Naatanen, 1992).

### ***Language-Specific Phoneme Representation in the Brain as Revealed by Means of the MMN***

Studies of categorical perception of speech have shown that speakers of different languages categorize speech sounds differently (Abramson & Lisker, 1967). The existence of language-dependent memory traces that can be activated in the presence of speech, but not when equally complex acoustic non-speech stimuli are processed, has been postulated (Kuhl, 1991, 1993). For example, Naatanen et al. (1997) recorded the MMN from Estonian and Finnish monolingual speakers using the vowels /e/, /ö/, and /õ/. Estonian has the three speech sounds in its vowel inventory while Finnish only has two (/e/ and /õ/) in his inventory. In that experiment, the second-formant (F2) of vowel /e/ was reduced in frequency by equal physical steps to generate a speech continuum ranging from /e/ to /õ/ passing through /ö/ (vowel /e/ and /õ/ having the highest and lowest F2 frequency, respectively). Vowel /e/ was used as a standard sound and vowels /ö/ and /õ/ served as deviant sounds.

The amplitude of the MMN was expected to increase as the acoustic difference between standard and deviant increased (Tiitinen, Reinikainen, & Naatanen, 1994). More specifically, the MMN waveform obtained from the difference between standard /e/ and deviant /ö/ was expected to be smaller than that obtained between standard /e/ and deviant /õ/. The results showed this amplitude pattern only occurred in Estonian speakers. That is, Estonians showed an increment in the MMN amplitude from deviant /ö/ to deviant /õ/ whereas Finnish speakers showed no amplitude differences from deviant

/ö/ to deviant /õ/. The lack of amplitude change between these deviant sounds suggested that Finnish speakers were not better discriminating /e/ from /ö/ than /e/ from /õ/. It was concluded that Finnish speakers are not sensitive enough to the acoustic properties of vowel /ö/ because those acoustic properties are not relevant in their language. This finding supported the idea of language-dependent memory traces that facilitates discrimination of native speech sounds (Kuhl, 1991, 1993).

The MMN has been helpful in investigating auditory processing and representation of phonemes in the brain. As mentioned earlier, the MMN is an indirect measure of auditory stimulus representation (Naatanen & Winkler, 1999). The MMN is believed to reflect the discrimination process between speech sounds and thus serves as an indirect measure of phonemic representations in the brain (Naatanen et al., 1997). In fact, the MMN literature supports the idea that exposure to one's native language deeply influences the way speech sounds are initially perceived (Best et al., 1988; Cheour et al., 1998; Kuhl et al., 1992; Naatanen et al., 1997; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005). That is, in the same way that monolingual speakers of different languages categorize speech differently, it has been shown that the MMN is language-dependent. Indeed, MMN amplitude differences resulting from language-specific memory traces have been widely reported for many languages (Nenonen, Shestakova, Huotilainen, & Naatanen, 2005; Peltola, Kujala, Tuomainen, Ek, Aaltonen, & Naatanen, 2003; Rivera-Gaxiola, Csibra, Johnson, & Karmiloff-Smith, 2000a; Rivera-Gaxiola, Johnson, Csibra, & Karmiloff-Smith, 2000b; Sharma & Dorman, 1998, 2000).

### *Sensory Stimulus Representation in Different Acoustic Contexts*

Little is known about how identical sounds presented in different acoustic contexts can be represented in the brain. Research with animals suggests that sound representation can be modified when acoustic signals are conditioned with specific behaviors. Studies have shown that cortical neural activity associated with particular acoustic features, such as frequency, can be modified by means of classical conditioning (Diamond & Weinberger, 1986; see Gao & Suga, 1998; Suga, Gao, Zhang, Ma, & Olsen, 2000 for mechanisms in sound representation). Yet, speech representation during different auditory contexts has not been widely explored. For example, it has not been determined whether language contexts can modify speech sound representation in the brain. The literature reviewed in this section focuses on the use of MMN (specifically the N2a component) to quantify changes in sound representation in the monolingual and bilingual brain.

Sussman, Winkler, Huotilainen, Ritter, and Naatanen (2002) asked whether the change-detection processes underlying the MMN could be affected by the instructions given to the participants. In their study Sussman and colleagues recorded the MMN from three groups of participants. There were three possible conditions in which the same stimuli were used, but each condition required a different task. The experiment consisted of one standard (440 Hz) and two deviant (494 Hz and 392 Hz) tones. The 494-Hz deviant tone was ignored in all conditions while the 392-Hz deviant tone was used to engage participants in one of three possible tasks. The tasks were as follow. The *ignore condition* consisted of participants ignoring all the stimuli. The *attend-pitch condition* consisted of actively identifying the 392-Hz tone. The last condition, the *attend-pattern*



*condition*, consisted of identifying a change in a tone-pattern. For example, participants heard the stimulus sequence 440, 440, 440, 440, 494 or 440, 440, 392, 440, 494 and were asked to press a button every time they heard a change in the sequence. In all conditions the MMN was based on tones that were not attended during the experiment (i.e., 440 Hz and 494 Hz). The results showed a robust N2a in the ignore condition and both an N2a and N2b response in the attend-pitch condition. The N2a and N2b brain responses were not obtained from the attend-pattern condition. The results implied that spectral differences between 440 Hz and 494 Hz were pre-attentively discriminated when tones were “processed” as independent spectral events, but the same tones were not discriminated when the task required integrating the spectral information into a stream of sounds.

The findings from the Sussman et al. (2002) experiment suggested that the MMN can be used as an electrophysiological tool to test bilinguals’ double phonetic standard. Specifically, it is possible that bilinguals represent speech sounds according to the language they are using at the moment. Winkler, Kujala, Alku, and Naatanen (2003) tested this idea by recording MMNs from Hungarian-Finnish bilinguals in two language contexts. In that study, the vowels /æ/ and /e/ were used because they represent different speech sounds in Finnish, but they are not discriminated in Hungarian (i.e., both vowels are perceived as /ε/). Specifically, the Finnish word *pæti* was used as the standard and the Finnish word *pēti* was used as the deviant. It was expected that bilinguals would perceive the deviant sound as *Pēti* (Peter) in the Hungarian language context and as *pēti* (bed) in the Finnish language context. Language contexts were maintained during ERP-recordings in two ways. First, participants were engaged in a conversation in the

language of interest before ERP recordings started. Second, during ERP recordings participants needed to press a response button every time they heard a target word in the language of interest. In this way, it was assumed that participants stayed focused on the language of interest. The data showed strong MMN (N2a and N2b/P2a complex) responses in both language contexts. The authors concluded that when a new phonetic distinction is learned (Hungarians learning Finnish /e/), the distinction is used irrespective of the language in use. These results imply that when L2 learners acquire a new phonetic sound that is not present in L1, the acoustic-to-phonemic representations of the new L2 speech sounds is not affected by language contexts.

In the Winkler et al. (2003) experiment, the Hungarian language context always preceded the Finnish language context. The investigators added an extra control condition to determine whether their results could have been biased by not counterbalancing the language contexts. The control condition consisted of recording the MMN in the Finnish language context during two consecutive sessions. The results showed that the mean amplitude between the MMN and following peak (i.e., N2b/P3a complex) was nearly significantly different ( $p = 0.08$ ), being greater in the first session than in the second session. Therefore, a language-context-related increase of the MMN amplitude (from the Hungarian to the Finnish language context) could have been masked by the fixed order of the two conditions.

In another investigation Garcia-Sierra and Champlin (2004) used stop consonants to assess pre-attentive discrimination in bilingual speakers of Spanish and English in two language contexts. In the first condition, bilinguals heard a sound with +15 ms of VOT and a /ga/ sound with -20 ms of VOT. Consonants with VOT durations of +15 ms

represents a /ka/ phonemic sound for Spanish speakers and a /ga/ phonemic sound for English speakers. In the second condition, bilinguals heard the same /ka/ sound and to another /ka/ sound with +50 ms of VOT. The /ka/ sound with +15 ms of VOT always functioned as deviant and the /ga/ or /ka/ (50 ms) functioned as standard sounds. Both conditions were presented in Spanish and English language contexts that were separated by 3 to 5 days.

Each pair of speech sounds was expected to be perceived as two sounds from the same phonetic category or as two sounds from different phonetic categories depending on the language context used. For example, when using the -20 ms VOT standard speech token, the deviant sound was expected to be perceived as a /ka/ sound in the Spanish language context and as /ga/ sound in the English language context. Similar results were expected when using the +50 ms VOT standard speech token. That is, the deviant sound was expected to be perceived as /ka/ in the Spanish language context and as /ga/ in the English language context.

The ERP recordings were divided into 5 intervals of 5-6 minutes each. During a given recording interval, participants read magazines in the language of interest and between recordings they were engaged in 3 minutes of conversations in the pertinent language. The results showed MMN amplitude did change in accordance with the language context under test. Specifically, bilinguals' showed greater MMN amplitudes when using the -20 ms VOT standard in the Spanish language context than in the English language context. Also, larger MMN amplitudes were found in the English language context than in the Spanish language context when using the +50 ms VOT standard sounds. These results suggested that in the case where a new phonetic distinction is not

learned (Spanish /g/ and English /k/ phonemes are similar), phoneme recognition can be influenced by available linguistic information.

In summary, the MMN can be used to assess the double phonemic boundary in bilingual. Still, the evidence in bilingual studies across language contexts is scarce and the MMN itself is not free from limitations.

### **Limitations in Assessing Phoneme Representation by Means of the MMN**

The limitations when assessing bilinguals' double phonemic boundary by means of electrophysiological techniques are, to some extent, the same as those affecting behavioral studies. Specifically, it has been questioned whether the MMN response truly reflects phonetic discrimination. When assessing categorical perception of speech, the MMN responses might reflect an acoustic rather than a phonetic distinction. In other words, if the mechanisms underlying the MMN reflect categorical perception of speech, then speech sounds belonging to the same phonetic category should not elicit an MMN, but speech sounds in a different category should elicit a MMN. The results obtained from speech-categorization studies have been inconclusive in this regard. For example, researchers using standard and deviant sounds drawn from the same phonetic category have reported the presence of an MMN response. This suggests that the MMN reflects acoustic changes rather than phonetic discrimination of speech (Rivera-Gaxiola et al., 2000a; Rivera-Gaxiola et al., 2000b; Sharma & Dorman, 1998).

Rivera-Gaxiola et al. (2000b) asked whether the MMN could be recorded from English monolingual speakers when they heard pairs of speech sounds taken from the same phonetic category. Four recordings were done, but only two are relevant for the purposes of the present investigation. The first recording was obtained with a standard

/ba/ that was an end-point on the speech continuum, and a deviant /ba/ close to the phonetic boundary. In the second recording, the same speech sounds were used, but the order of the stimuli was inverted (i.e., end point followed by boundary).

The results revealed an MMN in both conditions. The presence of the MMN in both conditions suggested that participants were able to differentiate both sounds regardless of the order of the stimuli. Similar results have been reported in Sharma and Dorman (1998). In Sharma and Dorman's experiment, an MMN was recorded using a 'good' representative (end point) as standard and a 'good' representative as deviant, and when using good and bad representatives (boundary) as standard and deviant, respectively. The investigators mentioned that constant repetition of the standard sound might favor brain electric activity associated with an acoustic response rather than with a phonetic response.

Phillips et al. (2000) designed a method that helped reduce the acoustic contrast between standard and deviant stimuli when recording the MMN. In the Phillips' et al. experiment, the MMN was recorded by delivering multiple standard and multiple deviant stimuli. This approach helps avoid a fixed acoustic difference between standard and deviant stimuli. Thus, when recording an MMN by means of multiple standards and deviants, the MMN response is less associated with an acoustic contrast, and perhaps more associated with phonemic distinctions. In the Phillips et al. experiment, a MMN was recorded when using speech sounds from different phonetic categories (/d/ and /t/), but no MMN was present when using standard and deviant sounds from the same phonetic category. Indeed, the stimulation sequence implemented when assessing speech categorization by means of MMN is an important variable to take into consideration. The

present investigation implemented 10 stimuli (5 standards and 5 deviants) to record the MMN. This technique attempts to minimize as much as possible a fixed acoustic difference between standard and deviant sounds.

#### **SUMMARY: LITERATURE REVIEW**

Studies assessing the double phonemic boundary in bilinguals have found different results. There is evidence suggesting that bilinguals have different phonemic boundaries from monolinguals, and that the phonemic boundaries do not shift back and forth in accordance to the language context (e.g., Caramazza et al., 1973; Williams, 1977). However, other researchers have suggested that the double phonemic boundary might emerge if bilinguals are kept focused on the language of interest throughout the entire experiment (e.g., using precursor sentences, see Elman et al., 1977). Indeed, using precursor sentences, other researchers have found that the phonemic boundary did shift back and forth depending on the language context (Flege & Eefting, 1987a; Hazan & Boulakia, 1993; Garcia-Sierra & Champlin, 2003). Bohn and Flege (1993) found that not only did bilinguals' voicing boundary shift, but also monolinguals' voicing boundary shifted in accordance with the language context used. These results suggested that the shift in the voicing boundary might be the consequence of biases caused by precursor sentences.

The Mismatch Negativity (MMN) has been used to test the idea that phonemes are represented in the brain (Naatanen et al., 1997) and that this representation is believed to be language specific (Best et al., 1988; Cheour et al., 1998; Kuhl et al., 1992; Naatanen et al., 1997; Rivera-Gaxiola et al., 2005). However, little is known about how identical sounds can be represented in the brain in different acoustic contexts. Although there is evidence suggesting that identical sounds (i.e., tones) can be discriminated in accordance

to the instructions given to the participants (Sussman et al., 2002), other evidence using speech sounds is contradictory. While Winkler et al. (2003) reported that identical speech sounds do not produce changes in the MMN in different language contexts, Garcia-Sierra and Champlin (2004) showed that the N2 deviant ERP response significantly changed according with the language context.

The first goal of this study was to test the double phonemic boundary controlling as much as possible for acoustic biases (i.e., precursor sentences). The second goal was to assess the MMN in bilinguals and monolinguals in two language contexts.

### **Chapter 3: Study Overview**

In the present study bilinguals' double phonetic representation was investigated. Specifically, phonetic decisions in two language contexts were assessed, and behavioral and electrophysiological measurements were obtained from monolingual speakers of English and from bilingual speakers of Spanish and English. Each of these steps is described below.

1. Phonetic decisions were indicated by having the listeners press a response button when perceiving the speech sound /t/ presented from a speech continuum varying from voiced to voiceless (/d/ to /t/) in 10 equal Voice Onset Time (VOT) steps. The 10 speech stimuli were presented in a random order.

2. Spanish and English language contexts were established by presenting videos and questions in the language of interest. Videos were presented before the identification task and the questions were presented during the identification task. The videos were used to prime participants in the language of interest. The questions were presented throughout the perceptual task and were used to keep participants engaged in the target language throughout the behavioral task.

3. Probability unit (PROBIT) analysis was used to determine the VOT value at 50% identification. PROBIT analysis associates a probability value with each stimulus based on the number of times each stimulus was presented and the number of times each stimulus was judged as /ta/. The VOT value associated with 0.5 probability was considered to be the phonemic boundary or 50% crossover. The 50% crossover point was derived for each participant in both language contexts.



4. Phonetic decisions were evaluated by means of signal detection theory (Macmillan & Creelman, 2005). Signal detection theory relates choice behavior to a psychological decision space. Choice behavior is quantified by the proportions of Hits and False-Alarm responses given to a set of stimuli. In the present study, pressing the response button when stimuli 5 to 25 ms of VOT were presented was defined as Hit responses, and pressing the response button when stimuli -25 to 0 ms of VOT were presented was counted as a False-Alarm response. The sensitivity measure ( $d'$ ) was calculated as the difference in Hit responses (in z-scores) between one stimulus and the next adjacent stimulus along the speech continuum (Burlingame, Sussman, Gillam, & Hay, 2005; Macmillan & Creelman, 2005). The formula that was used to calculate  $d'$  was:  $d' = z(\% \text{ identification item } x) - (\% \text{ identification } x + 1)$ .

Figure 2 shows the  $d'$  values and percent correct responses obtained from a hypothetical bilingual participant in both language contexts. The top panel shows cumulative  $d'$  values plotted against stimuli that are equally spaced in physical units (VOT). The graph compares the physical and psychological spacing of the stimuli, and its slope reveals how rapidly the perceptual effect grows with stimulus value. That is, it shows how sensitive the bilingual is to systematic stimulus changes in both language contexts. The bottom panel in Figure 2 shows that deviant sounds were judged more times as /t/ sound in the Spanish language context than in the English language context. The value of cumulative  $d'$  values obtained between both endpoint stimuli (-20 and +25 ms of VOT) represent the total sensitivity of the bilingual to the stimulus set and is called total  $d'$ . In the example plotted in Figure 2, total  $d'$  value was 5.22 in the Spanish language context and 4.92 in the English language context.

5. Event Related Potentials (ERPs) were measured concurrently with the behavioral task. Individual ERPs were averaged in two buffers while participants gave their responses to the speech stimuli. One buffer contained ERPs associated with standard or non-target sounds (/d/) and another buffer contained ERPs associated with the infrequent or deviant sound (/t/). The amplitude of the ERPs obtained from the standard and deviant sounds and the resulting MMN were analyzed.

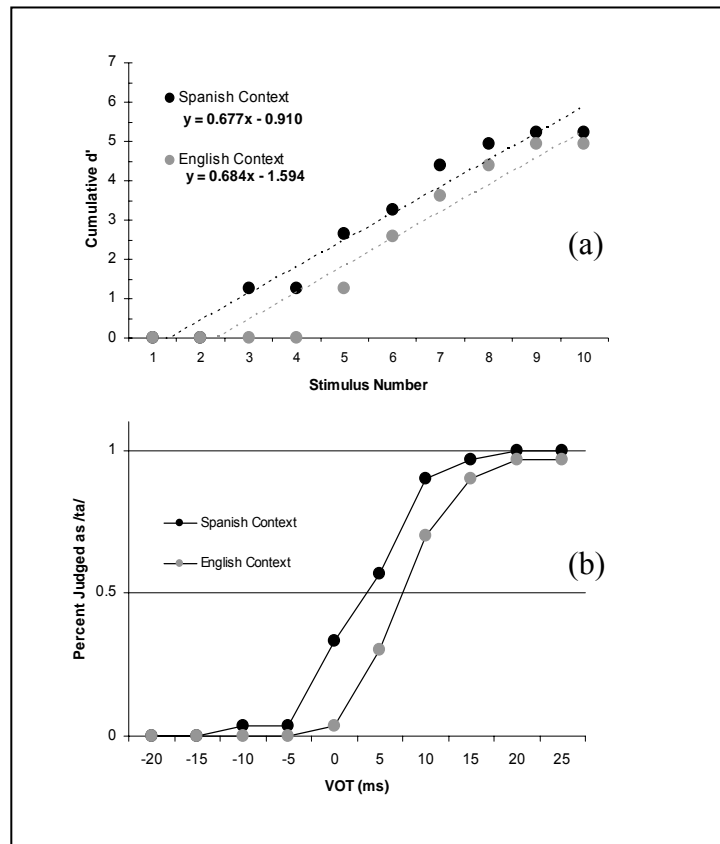


Figure 2: Hypothetical example of cumulative  $d'$  values in two language contexts. Note: (a) Cumulative  $d'$  values, and (b) percent “ta” responses obtained from a bilingual in two language contexts.

Figure 3 shows the number of times each stimulus was presented as a function of VOT. Standard sounds were presented with a probability of 0.8 and deviant sounds with

a probability of 0.2. The speech continuum consisted of sounds ranging from -20 to +25 ms of VOT in 5-ms steps (varying perceptually from /d/ to /t/). The stimuli with VOT values from -20 to 0 ms were used as standard /da/ sounds and the speech sounds ranging from 5 to +25 ms of VOT were used as deviant /ta/ sounds. These deviant sounds are likely to be perceived as /t/ sounds by monolingual speakers of Spanish, whereas they are likely to be perceived as /d/ sounds by monolingual speakers of English. Therefore, it was expected that bilinguals would perceive the deviants sounds according to the language context in use.

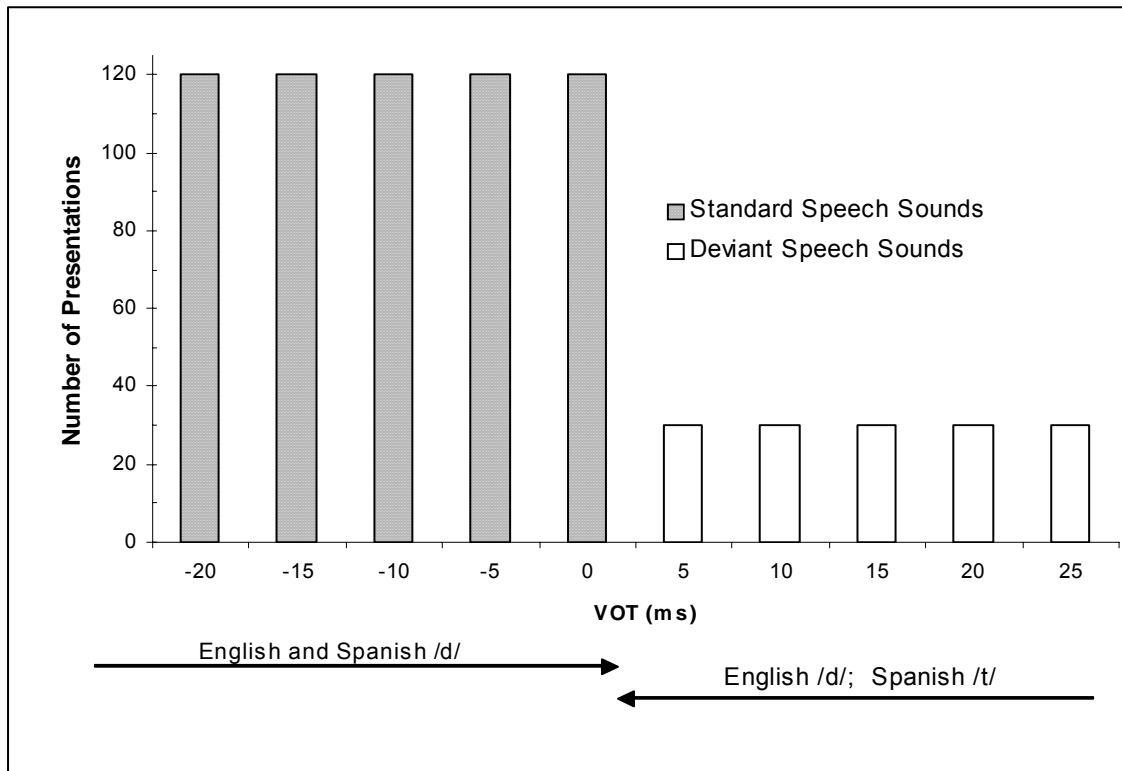


Figure 3: VOT values from which phonetic decisions and ERP responses were obtained in this study.

6. The ERP components P2, N2, and MMN were measured. In accordance with Hillyard, Hink, Schwent, and Picton (1973), the ERP component amplitudes were measured from baseline to peak. Baseline was defined as the mean voltage over the pre-

stimulus interval (0 to 100 ms). The peaks from P2 and N2 ERP components were selected as the most positive and most negative peak, respectively, in the time range from 150 to 350 ms after stimulus onset. Also, the standard ERP waveform was subtracted from the deviant ERP waveform in order to obtain the MMN. The MMN was calculated as the most negative peak in the time range from 200-298 ms after stimulus onset.

## **RESEARCH QUESTIONS**

### **Do bilinguals categorize speech sounds differently depending on the language context?**

Of interest in the present study was whether double phonetic representation occurs in bilinguals or is it the consequence of methodological bias. In order to test this idea, monolinguals and bilinguals were tested in the English and Spanish language contexts. It was hypothesized that bilinguals would show a different proportion of Hits (/t/) between language contexts and that monolinguals would show no or little change in the Hit proportion between language contexts.

### **Does the amplitude of the MMN in bilinguals depend on the language context?**

In order to test this question, the ERP responses of both bilinguals and monolinguals were measured. It has been shown that the amplitude of the MMN increases as the ability to detect acoustic differences between standard and deviant increases (Tiitinen et al., 1994). Therefore, it was hypothesized that bilinguals' P2, N2, and MMN ERP components would show larger amplitudes in the Spanish language context than in the English language context. Specifically, it was expected that bilinguals' phonemic boundary would shift toward the voiced endpoint in the Spanish language context, and toward the voiceless endpoint in the English language context.

The shift of the /t/-phonemic boundary toward the voiced endpoint would result in the perception of more 't' sounds than 'd' sounds, and thus, bilinguals would likely discriminate more 't' sounds in the Spanish language context than in the English language context. It was expected that the increased perception of 't' sounds in the Spanish language context would produce larger amplitude in the ERP components in bilinguals, but not in monolinguals.

Figure 4 shows the ERP responses from a bilingual in a pilot study in both language contexts. The P2 and N2 ERP components and the MMN showed a change in amplitude between language contexts. The MMN showed an amplitude of -3.8  $\mu\text{V}$  in the English language context and an amplitude of -4.87  $\mu\text{V}$  in the Spanish language context.

#### **Is the shift in the voicing boundary correlated with physiological measures?**

It was hypothesized that the shift in the bilinguals' voicing boundary would correlate with ERP amplitudes. Specifically, it was expected that bilinguals' phonemic boundary would shift toward the voiced endpoint in the Spanish language context, and toward the voiceless endpoint in the English language context. It was expected that shifts in the voicing boundary toward the voiceless endpoint would correlate with small ERP amplitudes, and shifts toward the voiced endpoint would correlate with larger ERP amplitudes. This expectation was hypothesized to occur in bilinguals, but not in monolinguals.

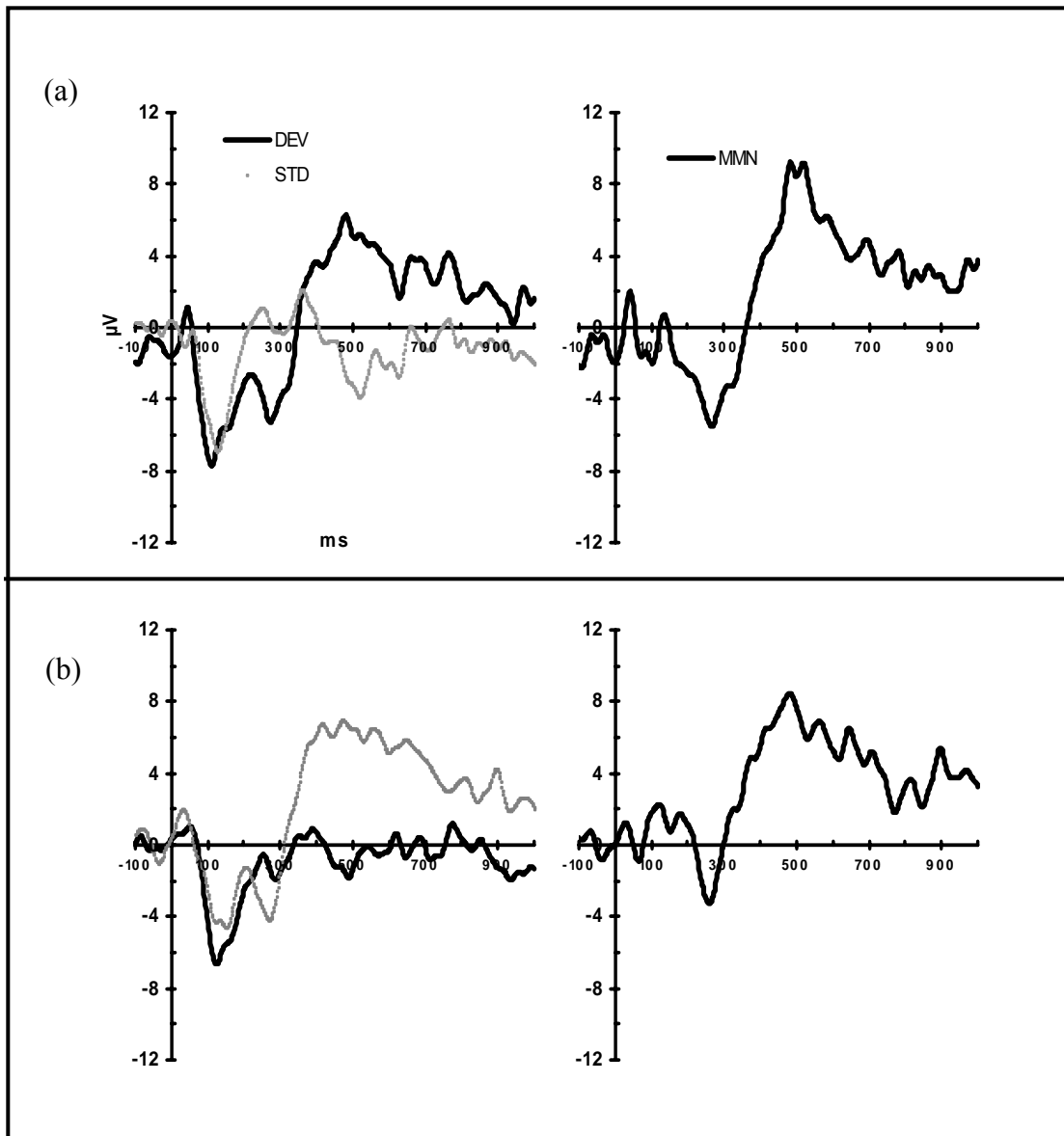


Figure 4: ERPs from a bilingual in both language contexts. Note: (a) ERPs collected in the Spanish language context; (b) ERPs collected in the English language context; DEV = Deviant Wave; STD = Standard Wave

## Chapter 4: Method

### PARTICIPANTS

Thirty one bilingual speakers of Spanish and English, and 32 monolingual speakers of English participated in the study. Only 54 participants were retained for subsequent analyses because they showed clear phonemic boundaries, clear ERP responses, and passed a hearing test. The final sample was 27 bilinguals (15 women and 12 men; mean age = 22.07;  $SD = 3.55$ ), and 27 monolinguals (13 women and 14 men; mean age = 22.55;  $SD = 3.69$ ). All participants were students from the University of Texas at Austin. Participants were recruited by means of flyers and they were paid \$30 for a 2 hour experimental session.

The target sample size for this study was calculated with the formula:

$$n = \frac{[Z_{\alpha/2} \times \sigma]^2}{E}$$

Where  $Z_{\alpha/2}$  is the z-score for alpha 0.05 (1.96), and  $\sigma$  (sigma) is the population standard deviation. In this case, because the population standard deviation is unknown, sigma was replaced with the sample standard deviation from previous research projects conducted by the experimenter and other investigators. The standard deviation obtained from phonemic boundaries (i.e., 50% cross over) from previous research projects with sample size larger than 30 was 10. Finally, E is the percent difference between population mean and the sample mean. This value is set by the experimenter and normally is 5 (i.e., .05).

Participants were defined as bilinguals or monolinguals by means of the language or languages they were exposed to throughout their lives. Bilingual candidates were considered only if they were exposed to both Spanish and English during childhood and

if they considered themselves to be fluent in both languages. Monolingual candidates were considered if they were exposed only to English during childhood, and if they did not consider themselves fluent in any other language than English. Participants who met the language requirements were then asked to answer a language questionnaire that assesses level of bilingualism (see Appendix A). Only those bilinguals who reported to be 75% (or above) confidence in reading, speaking, and listening in Spanish and English were invited to participate in the experiment. Monolinguals, on the other hand, were invited only if they reported to be 25% (or less) confident in reading, speaking, and listening in Spanish.

### **Bilinguals Usage of English and Spanish**

Fourteen bilingual participants were born in the U.S. and 12 participants were born in Spanish-speaking countries (i.e., Mexico = 7, Chile = 4, Uruguay = 1). Bilingual born in Spanish-speaking country reported to have been living in the U.S. an average of 15.12 years ( $SD = 8.00$ ). Sixty percent of the participants identified themselves with both the American and Mexican culture. Most participants (85%) used English 100% of the time to talk to employers and teachers, and most participants (77%) used Spanish 90%-100% of the time to talk to their mother or father. This information indicates that bilinguals used both English and Spanish in their daily lives to communicate with others.

Regarding self-reported mean confidence for *speaking* English and Spanish, a Likert scale from 1 to 5 (1 = “I cannot speak the language, I have a few words or phrases and, I cannot produce sentences,” and 5 = “I have a native-like proficiency with few grammatical errors and I have good vocabulary”) revealed a mean of 4.72 ( $SD = .46$ ) for both English and Spanish. Regarding self-reported mean confidence for *understanding*



English and Spanish, a Likert scale from 1 to 5 (1 = “I only understand a few words of what is being said,” and 5 = “I understand all of what is being said”) revealed a mean of 5.00 ( $SD = .00$ ) for English and a mean of 4.84 ( $SD = .37$ ) for Spanish. These findings indicate that bilinguals were proficient in speaking and understanding both English and Spanish.

## **QUESTIONNAIRES**

Two questionnaires were used in this study. They were a language background questionnaire and a Big Five Inventory in English and Spanish, which focuses on personality traits. All participants provided answers to these questionnaires. Each questionnaire is described next.

### **Language background questionnaire**

This questionnaire consists of 3 sections. The first section measures the *exposure* of English and/or Spanish. This section asks which language participants heard, spoke, and read the most from 3 years old to present time. The second section assesses the person’s *confidence* using each of these languages. Specifically, this section assesses the level of understanding, speaking, and reading Spanish and/or English at present and from 3 years old to present time. The last section asks the percentage of English and Spanish usage in regular day activities at present time (see Appendix A).

### **Big Five Inventory (BFI)**

The Big Five Inventory (BFI) is a questionnaire that measures five dimensions of personality. Both the BFI questionnaire in English (John & Srivastava, 1999) and in Spanish (Benet-Martínez & John, 1998) have 44 items with a 5-point-Likert scale, that ranges from 1 (*disagree strongly*) to 5 (*agree strongly*). This questionnaire was used in

this study to keep participants thinking in the language of interest. Only 18 questions were selected from this questionnaire and participants answered the questions during the experiment. Also the questions had an option with a question mark (i.e., meaning “I don’t know”) because it was expected that monolinguals would not understand the questions in Spanish, so they would use this option to answer each question.

### **GENERAL PROCEDURE**

The same general procedure was used for monolinguals and bilinguals. Potential participants were asked to participate in one experimental session of approximately 120 minutes. Before the start of the experiment, participants’ hearing was evaluated by means of a hearing screening in both ears. Participants with auditory thresholds that exceeded 20 dB at any frequency .25, .5, 1, 2, 4, 6, and 8 kHz, received \$5.00, and were excused from the experiment.

The experimental session consisted of assessing behavioral responses and brain electrical potentials in two language contexts. Both measurements, behavioral and electrophysiological, were done concurrently to control for participant state as much as possible throughout the experimental session (Hillyard, Squires, Bauer, & Lindsay, 1971).

### **Language Contexts**

Language contexts were established in two ways. 1) video-clips in the language of interest were shown on a computer monitor, and 2) questions in the language of interest were delivered throughout data collection. Participants watched a video-clip (4 min/each) before the start of a given language context. The video-clip was in English or Spanish depending on the language context. Video-clips were audible and subtitles in the

pertinent language were presented to facilitate participants' understanding of the events taking place in the video-clip. Participants answered four questions about specific events occurring in the video immediately after it ended. All questions were presented simultaneously through headphones and a computer monitor. Participants had approximately 12 seconds to answer each question.

After answering the questions, participants heard a pre-recorded male voice in the language of interest instructing them to get ready for the measurement phase. During the ERP and behavioral measurement phases, participants were instructed to press the response button every time they heard the speech sound /ta/. In order to avoid participant fatigue, a pause of 60s was inserted every 75s. During the pause, participants had time to relax and they also answered 2 questions in the language of interest (i.e., the selected questions from the personality questionnaire). These questions were delivered via headphones and appeared on the computer monitor. A total of 18 personality questions were given in the English condition and the same personality questions were delivered in the Spanish condition. Participants actively answered the questions by writing answers on a sheet of paper.

### **Stimuli**

Ten synthetic speech stimuli varying in VOT were generated using the cascade method described by Klatt (1980). All speech stimuli were 210 ms in duration with a 10-ms burst, 30-ms formant transition and 115 ms of steady-state (vowel). Stimuli were delivered at a rate of 1/s. The inter-stimulus time (ISI) varied from 1 to 1.2 s randomly. The stimuli were synthesized with five formants starting at appropriate /d/ onset frequencies values (i.e., F1 = 220 Hz; F2 = 1800 Hz; F3 = 3000 Hz; F4 = 3600 and F5 =

4500 Hz), and ramped to suitable /a/ frequency values ( $F1 = 720$  Hz;  $F2 = 1200$  Hz;  $F3 = 2770$  Hz;  $F4 = 3600$  and  $F5 = 4500$  Hz). Formant transitions were interpolated linearly over a time range of 30 ms.

English changes in voicing (voiced to voiceless) were accomplished by 1) delaying the energy in  $F1$  relative to the onset of higher formants and 2) by applying a noise source in  $F2$  and  $F3$  (amplitude of aspiration or  $AH = 65$ ) during the  $F1$  cutback period. Pre-voicing was accomplished by manipulating three parameters; fundamental frequency ( $F0$ ) amplitude of voicing ( $AV$ ) and amplitude of voice-excited parallel  $F1$  ( $A1V$ ) (Flege & Eefting, 1987).  $F0$  was set to 85 Hz,  $A1$  to 55 dB, and  $A1V$  to 45 dB throughout the pre-voicing period. An insert earphone (EAR Tone, model 3A 10 k $\Omega$ ) was used to present the speech sounds. The peak sound intensity (dB SPL) of each stimulus was measured with a sound-level meter that was connected to a 2-cc coupler. All stimuli were delivered at 85 dB peak-equivalent SPL, which is considered a comfortable listening level.

#### **BEHAVIORAL AND ELECTROPHYSIOLOGICAL PROCEDURE**

Participants' behavioral and electrophysiological responses were measured in two language contexts. Each language context included the same speech continuum to assess participants' behavioral and electrophysiological responses. In total, the stimuli were presented 750 times in a random order in each language context. The standard sounds occurred with a probability of 0.8 (600 stimulus repetitions) and the deviant sounds occurred with a probability of 0.2 (150 stimulus repetitions) (see Figure 3).

A language-context session lasted approximately 25 min. Each session consisted of 10 data collection phases (recording blocks) and 10 relaxation phases (relaxation

blocks) of approximately 75s/each. Each recording block had 75 speech stimuli (60 standards and 15 deviants). Within each block, the standard-to-deviant proportion was fixed; that is, the 75 speech sounds were distributed as 80% standard sounds and 20% deviant sounds. To be precise, within each recording block each standard sound was delivered 12 times and each deviant sound was delivered 3 times. The presentation of the 75 stimuli within each recording block was randomized, but 5 rules were considered in each pseudo-random sequence. 1) Deviants could not occur two times consecutively. 2) Between the offset of a deviant sound and the onset of the next deviant sounds, there must be at least three standard sounds. 3) The same standard sound could not occur two times consecutively. 4) Between the offset and the onset of a standard sound with the same VOT value, there must be at least three standard sounds with different VOT values. 5) In the entire stimulus sequence, each deviant sound was preceded 6 times by each standard sound. (see Appendix B for the stimulus sequence used in the 10 recording blocks).

### **Behavioral Procedure**

Participants were instructed to press the response button when hearing the syllable /ta/. Starting at the onset of the signal, participants had 1s to make a phonetic judgment and press the response button. Pressing the response button when stimuli from 5 to 25 ms of VOT were presented was scored as a Hit response, and pressing the response button when stimuli from -25 to 0 ms of VOT were presented was counted as a False-Alarm responses. Missed responses were calculated by subtracting Hits from the total number of deviant sounds presented. Correct Rejections were calculated by subtracting False-Alarm responses from total number of standard sounds presented.

## **Electrophysiological Procedure**

The electroencephalogram was recorded with gold-plated surface electrodes, NeuroScan SynAmp amplifiers, and Scan software from 6 inverting electrodes (Cz, Fz, Fp1, Fp2, M1, M2), one non-inverting electrode (tip of the nose), and one ground electrode (Fpz). All leads were placed according to the 10-20 International System. The M1 and M2 electrodes were used to assess MMN polarity inversion at the supra-temporal auditory cortex. Eye blinks were monitored with Fp1 and Fp2 electrodes. The recorded electroencephalogram was digitized at 500-Hz sampling rate and filtered using a band-pass filter with low and high cut-off frequencies at 0.05 Hz and 40 Hz, respectively. Epochs of 1000 ms with a 100 ms pre-stimulus interval were derived from the continuous electroencephalographic recording after off-line filtering the data with a band-pass filter from 0.1 to 30 Hz. Epochs with voltage changes exceeding +100 mV were omitted from the final average. The finale ERP waveforms were filtered using a band pass from 1.5 to 30 Hz (Roll-off of -12 dB/octave).

The time interval between the questions from the personality questionnaire and the first speech stimulus in any recording block was of 10s. This procedure ensured that the spectral information in the questions did not influence behavioral and/or electrophysiological responses (contrast effects) (Holt & Lotto, 2002; Lotto et al., 2003). Also, because contrast effects exert their effect mostly in speech sounds with ambiguous category membership (Diehl, Elman, & McCusker, 1978; Eimas & Corbit, 1973), the first speech token in a given recording block was always a standard sound with clear phonetic membership.

## **Chapter 5: Results**

The goal of this study was to assess the phonemic boundary obtained from a voiced to voiceless speech continuum in two language contexts. In order to accomplish this aim, behavioral and electrophysiological measures were obtained from bilingual speakers of Spanish and English and monolingual speakers of English.

### **BEHAVIORAL RESPONSES IN BILINGUALS AND MONOLINGUALS**

PROBIT analysis and Signal Detection Theory (SDT) were used to quantify the shift in the phonemic boundary in bilinguals and monolinguals in two language contexts. It was expected that bilinguals would shift the phonemic boundary toward the end of the voiced continuum in the Spanish language context and toward the end of the voiceless continuum in the English language context. Monolingual speakers of English, on the other hand, were not expected to show a shift in the phonemic boundary across language contexts.

### **Stimulus Selection**

The speech sounds used to assess the phonemic boundary were divided into deviant and standard stimuli (see Method section). The standard-to-deviant proportion was 1 deviant per 4 standards. The standard sounds that did not precede deviant sounds were not included in the final average. This procedure allowed including the same number of deviant and standard responses in the statistical analyses.

### **Probability Unit Analysis (PROBIT)**

Each participant's /ta/ cumulative distribution was transformed into a cumulative probability distribution by means of a PROBIT function. PROBIT analysis associates a probability value to each stimulus based in the number of times each stimulus was

presented, and the number of times each stimulus was judged /ta/. PROBIT was used to obtain phonemic 50% crossover in both language contexts.

Figure 5 shows how monolinguals and bilinguals categorized the stimuli as a function of VOT. The figure shows the percent (squares) and the probability (line) that each stimulus was perceived as 'ta'.

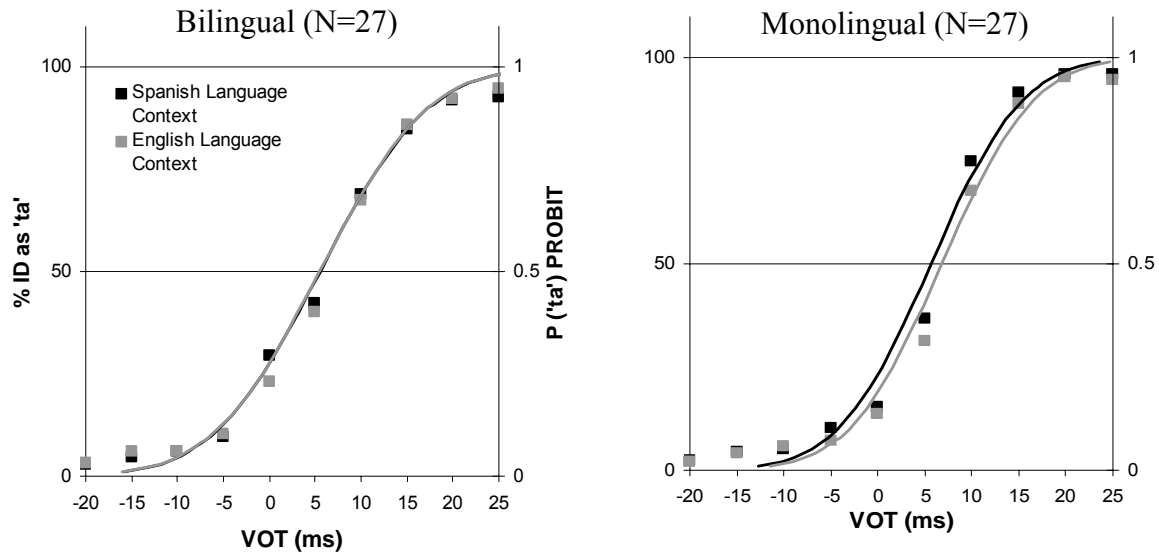


Figure 5: Bilinguals' and monolinguals' phonemic boundaries (50% point) in two language contexts derived from curves fitted via PROBIT analysis. Note: percent values (left y-axis) are indicated with squares and probability values (right y-axis) are indicated with lines.

Bilinguals showed that the phonemic boundary dividing the speech continuum did not change between language contexts. The phonemic boundaries from monolinguals showed a small effect in accordance with the language contexts. The VOT values at the 50% crossover point were compared in a 2 x 2 (group x language context) analysis of variance with repeated measures (language context). The ANOVA results showed that there were no main effects for language or group ( $F(1, 52) = 1.739, p = .193$ ,  $F(1, 52) = .295, p = .589$ ) and there was no significant interaction between language and group ( $F(1, 52) = 2.105, p = .153$ ).



## **Signal Detection Theory**

Signal detection theory relates a listener's choice behavior to a psychological decision space in which choice behavior is quantified by the proportions of Hit and False-Alarm responses given to the set of stimuli. SDT was used because the proportions of target sounds (/t/) and non-target sounds were not the same. In regular identification tasks, the stimuli are presented the same number of times (e.g., 10 times each). In the present investigation, brain electric responses were measured simultaneously with behavior, and therefore, it was necessary to change the basic identification paradigm into an odd-ball paradigm. The sensitivity measure  $d'$  (dee-prime) was used to verify that changes in the phonemic boundary were the consequence of participants being more sensitive to stimuli with +VOT (/t/) than to stimulus with –VOT (/d/). In the next section, results for the proportion of Hits and False-Alarms and the sensitivity measure  $d'$  are presented.

### ***Hit and False-Alarm Proportions***

The Hit and False-Alarm proportions were compiled to quantify the perception of /ta/. In accordance with SDT (Macmillan & Creelman, 2005), hit responses were defined as the number of times participants pressed the response button when target (deviant) sounds were presented. On the other hand, False-Alarm responses were defined as the number of times participants pressed the response button when non-target (standard) sounds were presented. It was expected that bilinguals, but not monolinguals would

show more Hit responses in the Spanish language context than in the English language context.

Figure 6 shows bilinguals and monolinguals Hit proportion as a function of VOT in both language contexts. It is important to note that in the figure, stimuli with 0, +5, and +10 ms of VOT were the ones that show the highest variability among language contexts in both groups. Thus, three independent 2 x 2 (group x language context) analysis of variance with repeated measures (language context) were performed for stimuli 0, +5, and +10 ms of VOT.

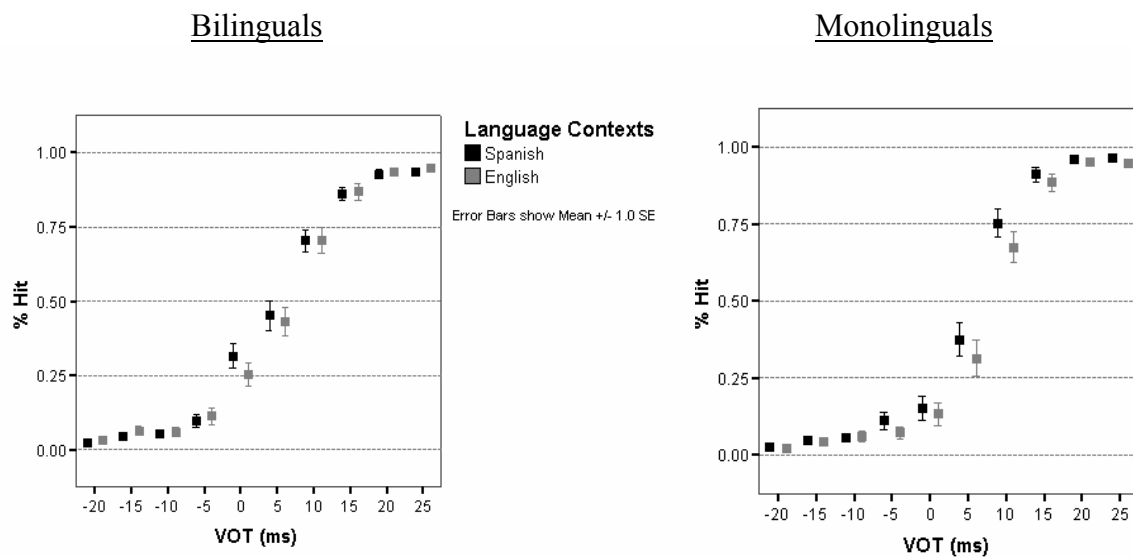


Figure 6: Bilinguals' and monolinguals' percent of hit responses in two language contexts as a function of VOT

The results for stimulus 0 ms of VOT showed a significant main effect for group and marginal significance for language ( $F(1, 52) = 8.20, p = .006$ , and  $F(1, 52) = 3.70, p = .061$ , respectively). There was no interaction between language and group ( $F(1, 52) = 1.01, p = .319$ ). Paired t-tests for stimulus 0 ms showed that bilinguals' mean Hit was larger in the Spanish context (Hit Mean = .32,  $SD = .21$ ) than in the English context (Hit Mean = .25,  $SD = .18, t(52) = 1.81, p = .08$ ). However, this difference did not reach

statistical significance. No differences were found for the monolinguals for stimulus 0 ms of VOT (Hit Mean in Spanish context = .15,  $SD = .16$ , and Hit Mean in English context = .13,  $SD = .19$ ,  $t(52) = .77$ ,  $p = .45$ ).

The results showed that bilinguals perception of stimulus 0 ms of VOT were in accordance to the prediction. That is, bilinguals will show more hit responses in the Spanish language context than in the English language context. Regarding stimulus +5 ms of VOT, the results showed no significant main effect for group or language ( $F(1, 52) = 1.88$ ,  $p = .175$ , and  $F(1, 52) = 2.37$ ,  $p = .130$ ; respectively). There was no interaction between language and group ( $F(1, 52) = .61$ ,  $p = .44$ ). Finally, stimulus +10 ms of VOT showed no significant main effect for group or main effect for language ( $F(1, 52) = .03$ ,  $p = .86$ , and  $F(1, 52) = 2.54$ ,  $p = .12$ ; respectively). There was no interaction between language and group ( $F(1, 52) = 2.72$ ,  $p = .10$ ). No further pair-wise comparisons were done for these stimuli because no significant main effects or interactions were found.

### ***The Overall Sensitivity to the Speech Continuum***

The sensitivity measure  $d'$  relates to an observer decision by taking into account the total number of stimuli presented and the total number of possible responses that can be given to the stimulus. The sensitivity measure  $d'$  was calculated as the difference of Hit responses (in z-scores) between one stimulus and the next adjacent stimulus along the speech continuum (Macmillan & Creelman, 2005). The formula that was used to calculate  $d'$  was:  $d' = z(\% \text{ identification item } x) - (\% \text{ identification } x + 1)$ .

Figure 7 shows how physical equally spaced stimuli were perceived as VOT increased. The x-axis represents the stimulus number and the y-axis gives the cumulative  $d'$ . Note that speech tokens are not labeled as a function of VOT. In this case, any given

stimulus number has a cumulative  $d'$  value that represents the sum of  $d'$  values from all smaller number adjacent stimuli. For example, the cumulative  $d'$  value from stimulus 3 represents the sum of  $d'$  values from stimulus 1 and 2. The total sensitivity of a participant to the stimulus set can be calculated as the difference between  $d'$  values obtained from both endpoint stimuli. The subtraction of both  $d'$  values results in an overall  $d'$ .

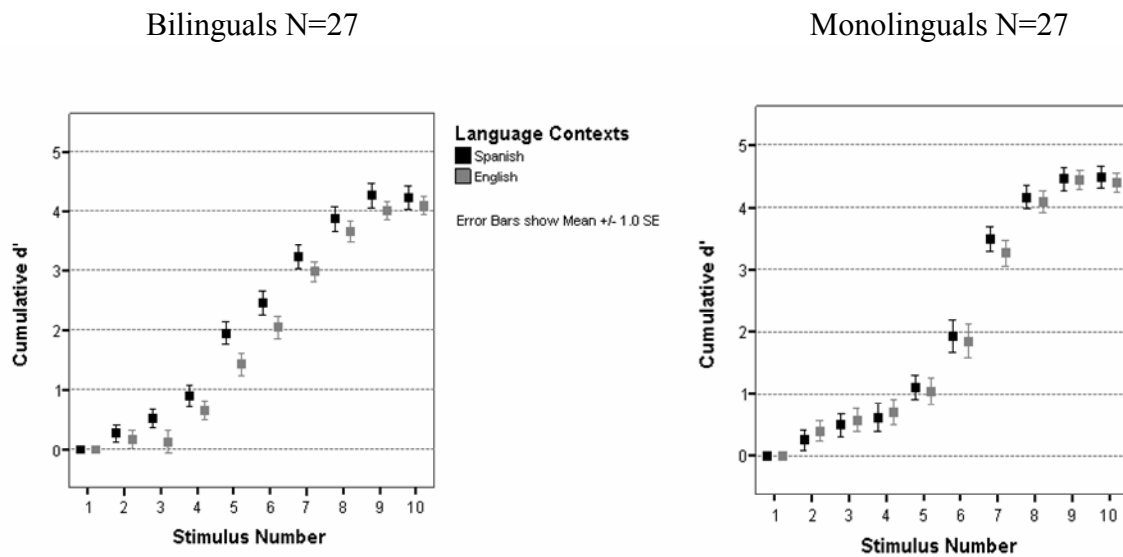


Figure 7: Bilinguals' and monolinguals' cumulative  $d'$  in two language contexts as a function of stimulus number.

Total  $d'$  was calculated in order to determine whether bilinguals have different sensitivities to the same set of stimuli depending on the language context. It was expected that bilinguals and not monolinguals would show different overall sensitivities between language contexts. For example, as it can be seen in Figure 7, the cumulative  $d'$  values go in the expected direction: bilinguals, but not monolinguals show larger sensitivities in the Spanish language context than in the English language context. In order to test this idea statistically, the total  $d'$  values were submitted to a separate 2 x 2 (group x language context) analysis of variance with repeated measures in the last factor.

The results showed no main effect for language or group ( $F(1, 52) = 1.281, p = .263$ ,  $F(1, 52) = .688, p = .411$ ; respectively), and there was no significant interaction between language and group ( $F(1, 52) = .519, p = .475$ ).

The results showed that the overall sensitivity of the speech continuum did not change among language contexts. However, visual inspection of Figure 7 shows that stimuli 5, 6, and 7 (i.e., parallels to stimuli 0, +5, and +10 ms of VOT, respectively) present marginal changes in the cumulative  $d'$  values among language contexts, especially in the bilinguals. This indicates that perhaps the cumulative  $d'$  values in this set of stimuli may have changed according to the language context. This idea was tested by conducting three independent  $2 \times 2$  (group  $\times$  language context) analysis of variance with repeated measures (language context) for stimuli 5, 6, and 7.

The results for stimulus number 5 showed a significant main effect for group and language ( $F(1, 52) = 7.00, p = .011$ , and  $F(1, 52) = 4.00, p = .05$ ; respectively). There was no significant interaction between group and language ( $F(1, 52) = 1.40, p = .12$ ). The results for stimulus number 6 showed no significant main effect for group or language ( $F(1, 52) = 1.70, p = .20$ , and  $F(1, 52) = 2.12, p = .15$ ; respectively). There was no interaction between language and group ( $F(1, 52) = .88, p = .35$ ). Finally, stimulus number 7 showed no significant main effect for group or language ( $F(1, 52) = 1.40, p = .24$ , and  $F(1, 52) = 2.57, p = .11$ ; respectively). There was no interaction between language and group ( $F(1, 52) = .011, p = .92$ ).

Table 2 shows the pair-wise comparisons across language contexts for each stimulus number for both bilinguals and monolinguals. The results indicated that there was a difference in stimulus number 5 and a marginal significance in stimulus 6 for

bilinguals. Monolinguals, on the other hand, did not differ across language contexts. Overall, bilinguals showed cumulative  $d'$  values that differed among language contexts. The cumulative  $d'$  values obtained in bilinguals in both language contexts were in accordance with the expectation. It was expected that bilinguals would show a larger cumulative  $d'$  values in the Spanish than in the English language context. This outcome suggests that bilinguals were more sensitive to systematic VOT changes in the Spanish language context than in the English language context.

Table 2: Bilinguals' and Monolinguals' Cumulative  $d'$  Means in Two Language Contexts

Bilinguals						
Spanish			English			
Stimulus No.	<i>Cum. d' Mean</i>	<i>SD</i>	<i>Cum. d' Mean</i>	<i>SD</i>	t-value	p-value
5	1.95	0.97	1.43	0.92	2.37	0.02
6	2.46	1.07	2.04	0.94	1.79	0.08
7	3.22	1.00	2.98	0.86	1.07	0.30
Monolinguals						
Spanish			English			
Stimulus No.	<i>Cum. d' Mean</i>	<i>SD</i>	<i>Cum. d' Mean</i>	<i>SD</i>	t-value	p-value
5	1.10	1.02	1.04	1.13	0.32	0.75
6	1.93	1.34	1.84	1.40	0.35	0.73
7	3.48	1.00	3.26	1.10	1.24	0.22

Note: Cum. = Cumulative, No. = Number

### Summary: Behavioral Analyses

The behavioral analyses can be separated into two approaches: 1) the PROBIT analysis utilized the 10 speech stimuli to quantify the degree to which the phonemic boundary shifted due to language contexts, and 2) Total  $d'$  quantified the response given to individual speech stimuli among the English and the Spanish language contexts. In general, when including the 10 stimuli into PROBIT or Total  $d'$  analysis, the results showed that bilinguals' and monolinguals responses did not change among language contexts. However, when only including stimuli 0, +5, +10 ms of VOT, the results

showed that bilinguals', but not monolinguals' judgments differed between language contexts.

## **ELECTROPHYSIOLOGICAL ANALYSES IN BILINGUALS AND MONOLINGUALS**

The electrophysiological results are presented in two sections. In the first section the analyses comparing the deviant response with respect to the standard response are shown. The purpose of this analysis is to quantify the presence of the MMN component in the deviant response. Specifically, the standard and deviant P2 and N2 components of the MMN are shown. In all cases amplitude analyses and graphical ERP waveforms are from the Cz electrode. The Cz electrode was chosen because it yielded the clearest response components. The second section analyses were conducted to determine whether ERPs differed between language context in bilinguals and monolinguals. Specifically, results for the P2 and N2 components from the deviant ERP are presented and then the MMN amplitude will be reported.

### **Amplitude Differences between Standard and Deviant ERP Components**

The MMN typically is manifested in the deviant ERP as an amplitude increase in P2 and/or N2 components (Naatanen, 1992). The subtraction of the standard response from the deviant response can give a clear picture of the time range when the MMN component is present. It was expected that both monolinguals and bilinguals would show larger N2 and/or P2 amplitudes in the deviant response relative to the standard response.

Figure 8 depicts standard and deviant ERPs for both language contexts (Panels A and B), and the subtraction of the standard ERP waveform from the deviant ERP waveform (Panels C and D). The ERP waveforms in figure 8 show that deviant and

standard amplitudes from components P2 and N2 were different. Also, figure 8 shows that the MMN was indeed the consequence of N2 and P2 deviant components having different amplitudes when compared with the P2 and N2 standard components.



### Bilinguals' ERP Responses (N=27)

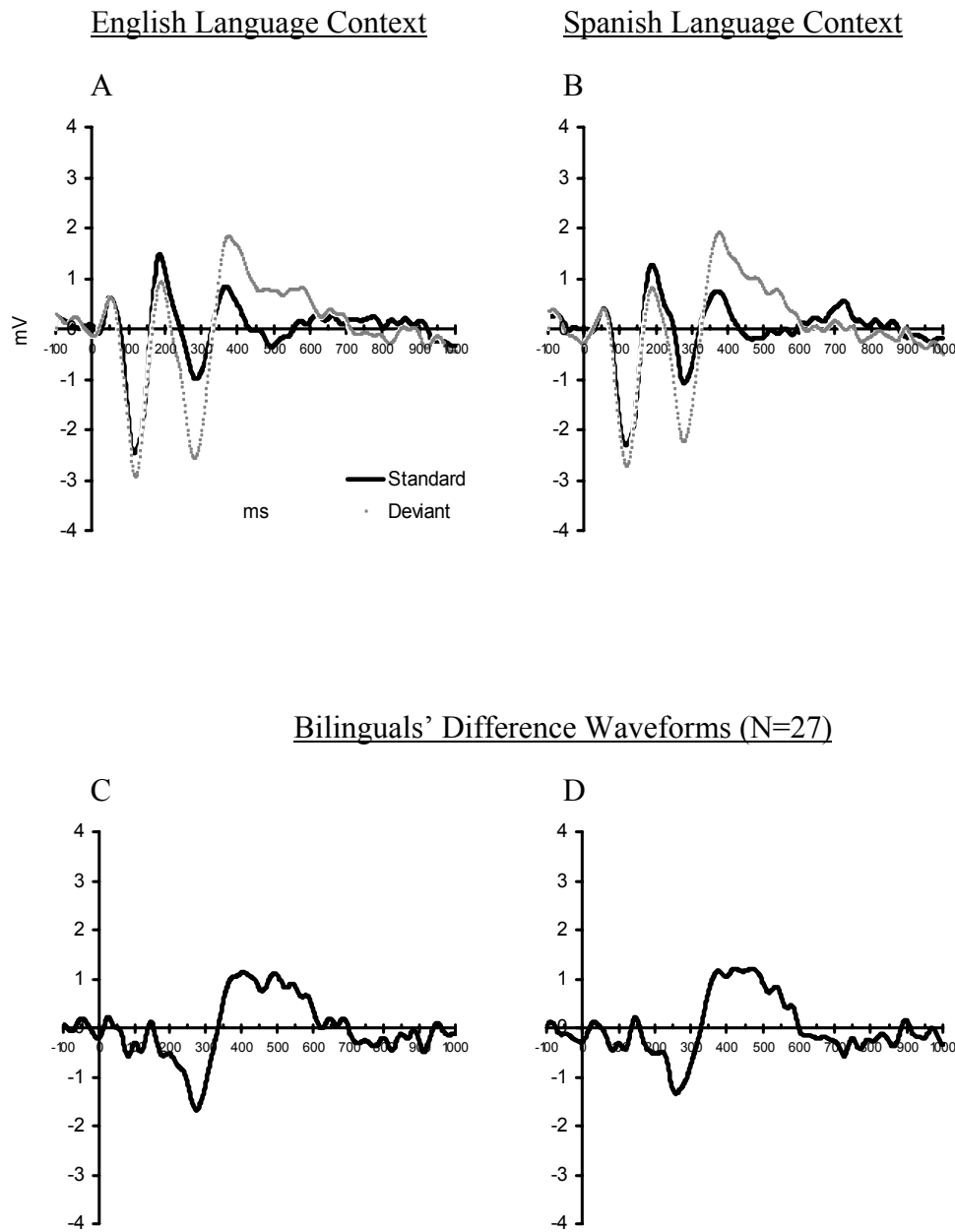


Figure 8: Bilinguals' ERP responses associated with standard and deviant sounds in two language contexts. Note: Panels A and B are the raw waveforms and Panels C and D are the difference waveforms obtained in the English and Spanish contexts, respectively.

### Monolinguals' ERP Responses (N=27)

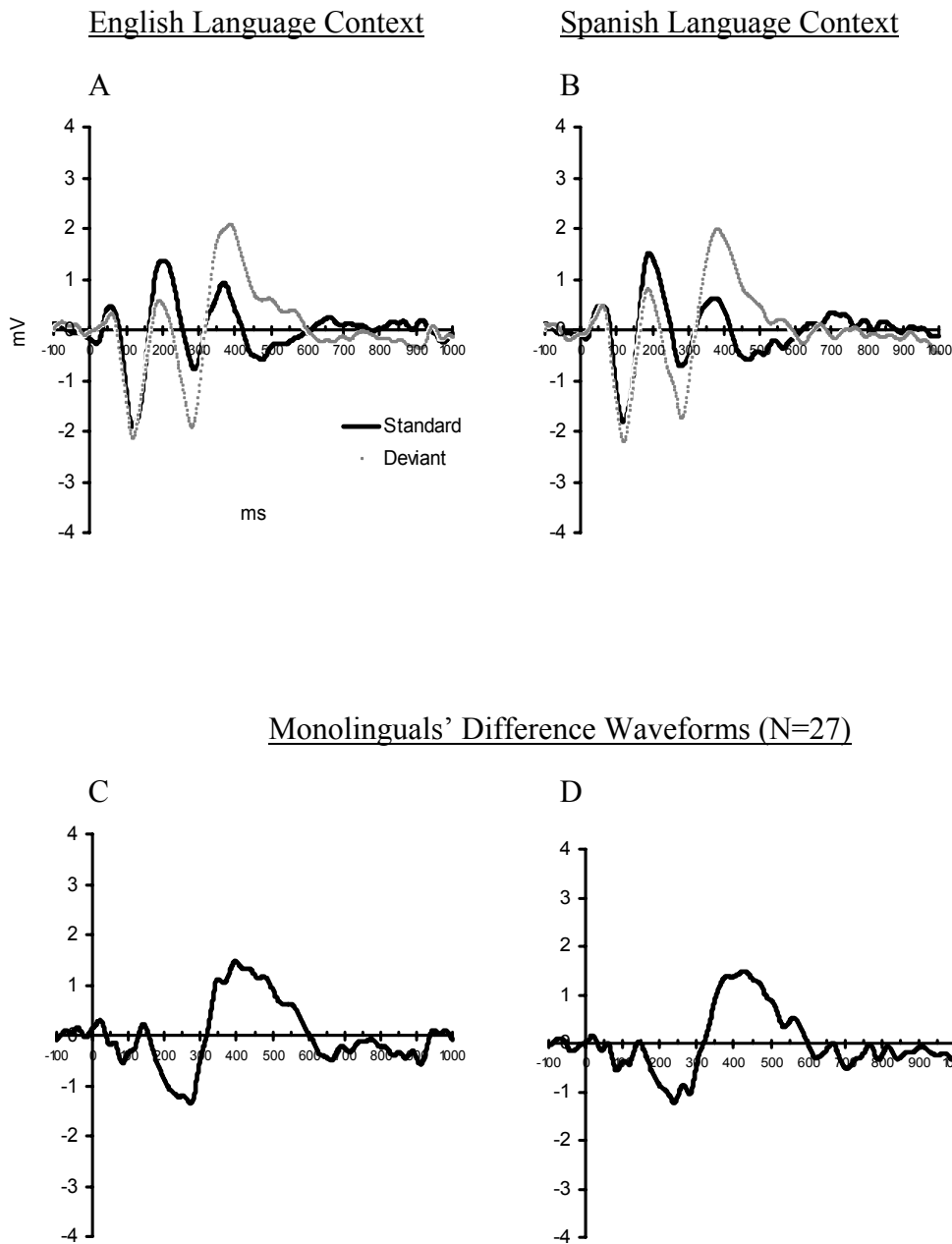


Figure 8: Continue: Monolinguals' ERP response associated with standard and deviant sounds in two language contexts. Panels A and B are the raw waveforms and Panels C and D are the difference waveforms obtained in the English and Spanish contexts, respectively.

The amplitudes (N2 and P2) were submitted to a separate 2 x 2 (group x stimulus type) analysis of variance with repeated measures on the last factor for both the English and Spanish language contexts. Results for the P2 component in the English language context showed no significant main effect for group ( $F(1, 52) = .13, p = .71$ ), but did show a significant main effect for stimulus type (i.e., standard vs. deviant) ( $F(1, 52) = 34.02, p < .001$ ); the group by stimulus type interaction was not significant ( $F(1, 52) = .34, p = .56$ ). The Spanish language contexts, P2 amplitude showed no significant main effect for group ( $F(1, 52) = .28, p = .60$ ), but did show a significant main effect for stimulus type (i.e., standard vs. deviant) ( $F(1, 52) = 9.60, p < .01$ ); the group by stimulus type interaction was not significant ( $F(1, 52) = .85, p = .36$ ).

Results for the N2 component in the English language context showed a significant main effect for group ( $F(1, 52) = 6.10, p < .05$ ), but did show a significant main effect for stimulus type (i.e., standard vs. deviant) ( $F(1, 52) = 44.66, p < .001$ ); the group by stimulus type interaction was not significant ( $F(1, 52) = .55, p = .46$ ). In the Spanish language contexts N2 amplitude showed no significant main effect for group ( $F(1, 52) = 1.98, p = .16$ ), but did show a significant main effect for stimulus type (i.e., standard vs. deviant) ( $F(1, 52) = 27.50, p < .001$ ); the group by stimulus type interaction was not significant ( $F(1, 52) = .56, p = .45$ ).

It is worth noting that in the English language context, there was a difference in the N2 amplitude for group. Specifically, the deviant for the bilinguals was significantly lower than the deviant from the monolinguals. Table 3 shows means and standard deviations for standard and deviant for each P2 and N2 component in bilinguals and monolinguals in two language contexts.

Table 3: Means and Standard Deviations for Standard and Deviant P2 and N2 ERP Components in Two Language Contexts

ERP Component								
<i>English Language Context</i>	P2				N2			
	Standard		Deviant		Standard		Deviant	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
Bilinguals	1.88	1.17	1.13	1.12	-1.82	1.4	-3.2	1.5
Monolinguals	2.07	1.31	1.16	1.17	-1.15	1.3	-2.25	1.32
<i>Spanish Language Context</i>	P2				N2			
	Standard		Deviant		Standard		Deviant	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
Bilinguals	1.72	0.98	1.37	1.46	-1.64	1.06	-2.7	1.52
Monolinguals	2.04	1.5	1.38	1.25	-1.36	1.1	-2.15	1.3

### Electrophysiological Measurements across Language Contexts

In this section the differences across language contexts for components P2 and N2 are presented, followed by the MMN results.

#### *N2 and P2 Deviant ERP Components*

It was expected that bilinguals, but not monolinguals, would show larger amplitudes for N2 and/or P2 ERP components in the Spanish language context than in the English language context. It was expected that bilinguals' phonemic boundary would shift toward the voiced endpoint in the Spanish language context, and toward the voiceless endpoint in the English language context. The shift of the /t/-phonemic boundary toward the voiced endpoint would cause the perception of more 't' sounds than 'd' sounds, and thus, bilinguals would likely discriminate more 't' sounds in the Spanish language context than in the English language context. It has been shown that the amplitude of the P2 and N2 components in the deviant ERP (i.e., MMN) increases as the ability to detect acoustic differences between standard and deviant increases (Tiitinen et al., 1994). Therefore, it was expected that bilinguals' increased ability to detect deviant

sounds from standards sounds in the Spanish language context would result in an increase in amplitude in the mentioned components from the deviant ERP.

Figure 9 shows bilinguals' and monolinguals' ERP components (i.e., deviant P2 and N2) in both language contexts. The figure depicts that N2, but not P2 amplitude is larger for bilinguals in the English context than in the Spanish context. Monolinguals, on the other hand, did not show the same trend.

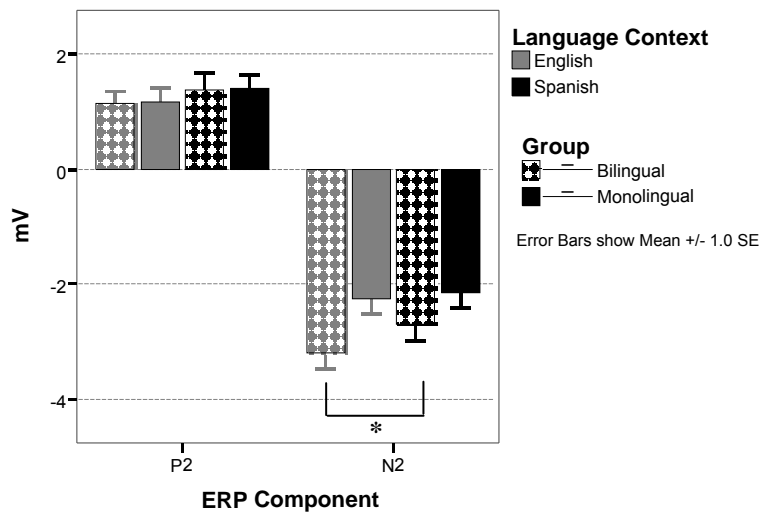


Figure 9: Bilinguals' and monolinguals' P2 and N2 deviant ERP components in both language contexts. Note: \*  $p = .047$ ; Error bars show Mean  $\pm$  1.0 SE.

The amplitudes of P2 and N2 components obtained from the deviant response were submitted to a separate 2 x 2 (group x language context) analysis of variance with repeated measures (language context). The results for the P2 component showed no significant main effect for group or language context, and the interaction between group

and language context was not significant ( $F(1, 52) = .005$ ,  $p = .946$ ,  $F(1, 52) = 3.58$ ,  $p = .064$ , and  $F(1, 52) = .001$ ,  $p = .971$ ).

With regard to the N2 component, there was a significant main effect for group ( $F(1, 52) = 4.65$ ,  $p < .05$ ), no significant main effect for language ( $F(1, 52) = 3.11$ ,  $p = .083$ ), and no significant interaction between group and language context ( $F(1, 52) = 1.382$ ,  $p = .245$ ). Independent t-test comparisons for group showed that bilinguals (Deviant Mean = -3.19,  $SD = 1.50$ ) and monolinguals (Deviant Mean = -2.25,  $SD = 1.32$ ) differed only in the English language context ( $t(52) = -2.45$ ,  $p = .018$ ). Because the main effect for language almost reached significance, paired t-tests were done. Results revealed that only bilinguals differed across language contexts (English Deviant Mean = -3.19,  $SD = 1.50$ , and Spanish Deviant Mean = 2.70,  $SD = 1.52$ ,  $t(52) = -2.10$ ,  $p = .047$ ).

### ***Mismatch Negativity (MMN)***

It was expected that bilinguals, but not monolinguals, would show larger MMN amplitudes in the Spanish language context than in the English language context. As it was described for the N2 deviant ERP component, the MMN amplitude was expected to increase in amplitude as the /t/-phonemic boundary shifted toward the voiced endpoint, and decrease in amplitude as the /t/-phonemic boundary shifted towards the voiceless endpoint. The MMN was calculated by choosing the most negative valley in the time range between 150-350 ms after stimulus onset. The most negative value was subtracted from baseline in order to estimate the amplitude. Figure 10 shows the MMN amplitudes for bilingual and monolinguals in both language contexts. Bilinguals, but not monolinguals showed a larger amplitude in the English language context than in the Spanish language context.

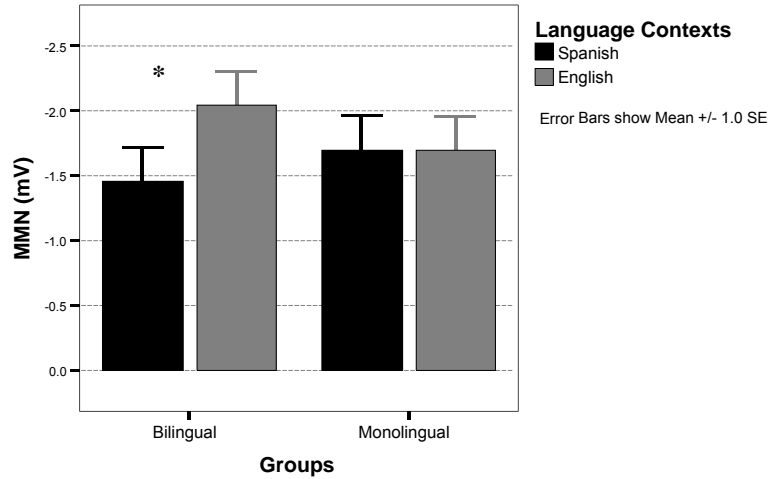


Figure 10: Bilinguals' and monolinguals' MMN amplitudes in both language contexts. Note \*  $p = .11$ ; Error Bars show Mean  $\pm$  1.0 SE.

The MMN amplitudes were submitted to a 2 x 2 (group x language context) analysis of variance with repeated measures (language context). The results showed no significant main effect for group, language context or interaction between group and language contexts ( $F(1, 52) = .034, p = .854$ ,  $F(1, 52) = 1.55, p = .219$ , and  $F(1, 52) = 1.57, p = .216$ ). The pair-wise comparison for bilinguals was close to significance but not for the monolinguals ( $t(26) = 1.67, p = .11$ ;  $t(26) = -.006, p = .99$ ; respectively).

Bilinguals' MMN amplitudes did follow the pattern of the results obtained from the N2 deviant ERP component. That is, contrary to expectation, the amplitude of the MMN was greater in the English language context than in the Spanish language context. However, in this case, the MMN amplitude did not reach significance ( $p = .11$ ). These results support the idea that bilinguals were better at detecting phonetic differences between standard and deviant in the English language context than in the Spanish language context.

## CORRELATIONS BETWEEN ELECTROPHYSIOLOGY AND BEHAVIOR

The third goal of the present investigation was to examine the relationship between the behavioral and electrophysiological responses. More specifically, bilinguals' double phonemic boundary was compared to the amplitude of the ERP components. In order to accomplish this aim, the VOT values obtained for the 50% crossover from the Spanish language context were subtracted from the 50% crossover values obtained in the English language context (i.e.,  $VOT_d$ ). Next, the P2 peak (i.e.,  $P2_p$ ) was subtracted from the N2 valley (i.e.,  $N2_v$ ) in order to obtain an overall amplitude difference (i.e.,  $P2_p-N2_v$ ). Then  $P2_p-N2_v$  values from the Spanish language context were subtracted from the values obtained in the English language context. Finally, the correlations between  $VOT_d$  and  $P2_p-N2_v$  for both monolinguals and bilinguals were calculated.

It was hypothesized that  $VOT_d$  would correlate with  $P2_p-N2_v$  for bilinguals but not for monolinguals. In other words, if bilinguals categorized the VOT continuum differently depending on the language context, then the ERPs would also show a difference among language contexts. Figure 11 shows the scatter plots for both groups. The bilingual scatter plot shows that large shift in  $VOT_d$  resulted in large amplitudes in deviant  $P2_p-N2_v$ . In the other hand, the monolingual scatter plot did not show the same trend.



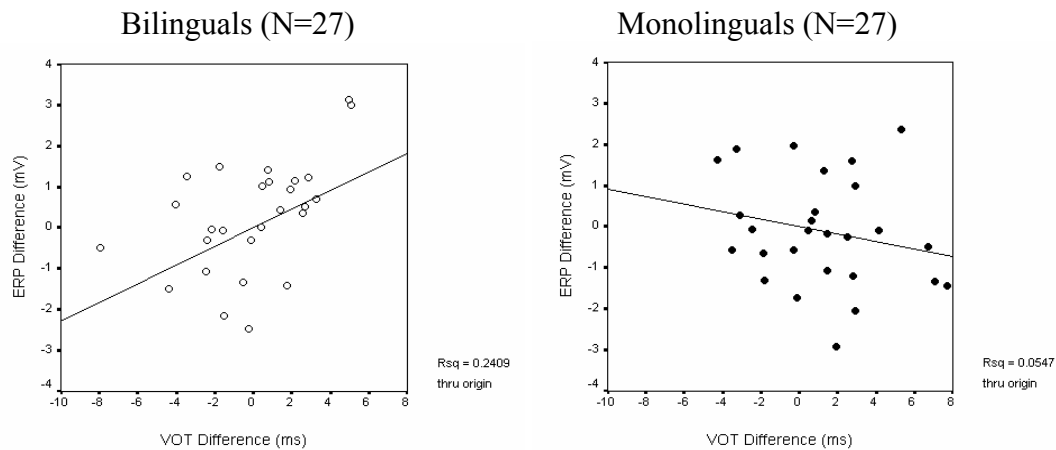


Figure 11: Behavior and ERP scatterplots. Note: ERP difference is plotted against VOT difference for bilinguals (left panel) and monolinguals (right panel)

Correlation analyses revealed that there was significant positive correlation for bilinguals ( $r = .504$ ,  $p < .01$ ), but not for monolinguals ( $r = -.21$ ,  $p = .282$ ). That is, large  $VOT_d$  correlated with large amplitude in the  $P2_p$ - $N2_v$  deviant components. This trend was found in bilinguals but not in monolinguals.

### Summary: Electrophysiological Analyses and Correlations

The results indicate that bilinguals indeed had a significant change in the amplitude of N2 deviant ERP component. However, the amplitude change was in the opposite direction to the one hypothesized. On the other hand, significant correlations were found for bilinguals, but not for monolinguals. Bilinguals had a positive correlation between the degree of shift of the voicing boundary and amplitude of  $P2_p$ - $N2_v$ . The correlations indicated that bilinguals' voicing boundary that shifted in the expected direction showed larger amplitudes. Whereas, bilinguals that did not show a shift or had a shift in the opposite direction presented smaller amplitudes.

## **Chapter 6: General Discussion**

One goal of this investigation was to determine whether bilinguals possess the double phonetic boundary. In order to accomplish this objective, bilingual and monolingual participants' behavioral and electrophysiological measures were taken in two language contexts. Specifically, bilingual speakers of Spanish and English, and monolingual speakers of English were asked to identify the speech sound /ta/ from a 10-token speech continuum ranging in VOT from /da/ to /ta/, while electrophysiological measures also were assessed concurrently using the same stimuli.

### **SPEECH CATEGORIZATION AND BILINGUALS' DOUBLE PHONEMIC BOUNDARY**

Past studies assessing the double phonemic boundary in bilinguals have found contradictory results. There is evidence suggesting that bilinguals have different phonemic boundaries from monolinguals and that their phonemic boundaries shift according to the language in use. For example, Caramazza et al. (1973) and Williams (1977) reported that bilinguals showed phonemic boundaries at different VOT values from monolinguals. Other researchers suggested that the double phonemic boundary emerges if bilinguals are kept focused on the language of interest throughout the entire experiment. Indeed, when researchers delivered precursor sentences in the language of interest throughout the perceptual task, bilinguals' perception of speech sounds close to the phonemic boundary (i.e., ambiguous sounds) was affected depending on the language in use (Elman et al., 1977).

Using precursor sentences, other researchers reported that the phonemic boundary dividing a voiced to voiceless VOT continuum shifted back and forth depending upon the language context (Flege & Eefting, 1987a; Hazan & Boulakia, 1993; Garcia-Sierra & Champlin, 2003). However, Bohn and Flege (1993), using the precursor-sentence

method, found that not only bilinguals' voicing boundary shifted, but also monolinguals' voicing boundary shifted according to the language context used. These results suggested that the shift in the voicing boundary was the consequence of biases caused by precursor sentences rather than by the existence of a double phonemic boundary in bilinguals.

One objective of this study was to test the double phonemic boundary by controlling for acoustic biases. Thus, the language contexts were established as follows. First, precursor sentences were not presented during the perceptual task, rather bilinguals were asked to watch videos and to answer questions in the language of interest. Videos were presented before the perceptual task and the questions were presented every 75s. There was a 10-s interstimulus interval between the last question delivered and the presentation of the first speech sound to be identified. Research has shown that acoustic information preceding phonetic judgment with interstimulus intervals longer than 1.3 s does not affect the placement of phonetic categories (Holt & Lotto, 2002; Lotto et al., 2003). Second, monolinguals and bilinguals were exposed to the same amount of Spanish and English before the perceptual task. Bilinguals were interviewed in Spanish at the end of the experiment to confirm that they were fluent in Spanish. Finally, similar to Bohn and Flege (1993), bilingual speakers of English and Spanish and monolingual speakers of English were tested in the Spanish and English language contexts. That is, monolinguals served as a control group.

The results of the present study showed that bilinguals' and monolinguals' double phonemic boundaries did not change in accordance with language contexts. This result is consistent with the findings of Caramazza et al. (1973) and Williams (1977) in the sense

that bilinguals did not show a double phonetic boundary. However, the present investigation also revealed that the phonemic boundaries from bilinguals and monolinguals did not differ.

It is important to mention that in this study, phonetic judgments were collected by means of an odd-ball paradigm. Normally, identification paradigms present each stimulus 10 times. In the present study, each standard sound was presented 120 times, and each deviant stimulus was presented 30 times in each language condition. This suggests that bilinguals and monolinguals might have learned, over time, the acoustic difference between speech tokens. As a result, both groups split the VOT continuum into two equal halves. Participants seemed to have assigned each stimulus into a category with -VOT or +VOT (crossover +5ms VOT approx.). This effect is known as a 'range effect' (Brady & Darwin, 1978; Keating, Mikos, & Ganong, 1981). Range effects are based on the assumption that, all things being equal, participants will prefer to split the range into two equal halves. That is, the perception of a particular speech token depended on the set of VOT stimuli, rather than on a particular linguistic category. This suggests that perhaps range effects influenced the final results regarding the shift of the boundary in accordance with the language.

Signal Detection Theory was used as a way to assess how physically equal spaced stimuli are perceived as VOT changes. Cumulative  $d'$  was calculated by obtaining the difference of Hit responses (in z-scores) between one stimulus and the next adjacent stimulus along the speech continuum (Macmillan & Creelman, 2005). The overall sensitivity (Total  $d'$ ) of the speech continuum obtained from bilinguals and monolinguals was not different between language contexts. However, bilinguals showed higher

sensitivities (cumulative  $d'$ ) in the Spanish language context than in the English language context for stimuli 5 and 6. The stimuli that showed a change in sensitivity were stimuli that were close or at the phonemic boundary.

Two possible reasons may explain the lack of shift in the bilingual voicing boundary. First, it could be that bilinguals do not have a double phonetic boundary. That is, considering that an effort was made to control for acoustic biases (i.e., carrier phrases were not used, acoustic information was interspaced by at least 10 seconds between stimuli, and language contexts were counterbalanced), it could be assumed that the established bilinguals' double phonemic boundary were a result of acoustic biases. Second, the odd-ball paradigm may not be an appropriate way to test the placement of phonemic boundaries in VOT continua. The fact that bilinguals, but not monolinguals, showed a small change in sensitivity (cumulative  $d'$ ) among language contexts suggests that hearing the same set of stimuli during 25m might have obscured bilinguals' double voicing boundary. In other words, the conditions of the experiment may not have allowed the double phonemic boundary to emerge.

In addition to the typical behavioral measures used to test the double phonetic boundary, electrophysiological measures were assessed. This physiological approach may provide some new insight regarding the bilinguals' double phonetic boundary.

#### **ELECTROPHYSIOLOGY AND BILINGUALS' DOUBLE PHONETIC BOUNDARY**

A second goal of the present investigation was to test the double phonemic boundary in bilinguals by measuring electrophysiological responses. Important to mention is the fact that this is the first investigation that attempted to tap the physiological mechanisms underlying bilinguals' double phonetic boundary. Mismatch

Negativity (MMN) has been used as an approach to test the idea that phonemes are represented in the brain and that this representation is language specific (Best et al., 1988; Cheour et al., 1998; Kuhl et al., 1992; Naatanen et al., 1997; Rivera-Gaxiola et al., 2005). MMN was used to test the idea that pre-attentive discrimination in bilinguals may vary in accordance with language contexts.

Little is known about how sounds are represented in the brain in different acoustic contexts. There is evidence suggesting that sounds (i.e., tones) can be discriminated in accordance to the instructions given to the participants (Sussman et al., 2002). There is also evidence that the discrimination of speech sounds does not produce changes in the MMN response when recorded in different language contexts. For example, Winkler et al. (2003) reported that Hungarian speakers of Finnish did not show a change in the MMN when discriminating vowel /e/ in a Hungarian language context and in a Finnish language context. These researchers concluded that when new phonetic contrasts are learned (Hungarians learning the vowel contrast /e/) the acoustic-to-phonemic representations of the new speech contrasts are not affected by language contexts.

On the other hand, Garcia-Sierra and Champlin (2004) used stop consonants to assess pre-attentive discrimination in bilingual speakers of Spanish and English in two language contexts. In their study, in contrast to the Winkler et al. (2003) experiment, the target (deviant = +15 ms VOT) speech sound that was used was shared among the English language and the Spanish language. Specifically, speech sounds with +15 ms of VOT represent a voiced phonetic category in English, but they represent a voiceless category in the Spanish language. The results showed that the N2 deviant ERP response significantly changed according to the language context. This suggests that speech

information can be represented differently depending on the language used at the moment.

The assessment of the MMN when using speech sounds has been criticized because it is hard to separate acoustic from phonetic discrimination. Specifically, when recording the MMN a repetitive sequence of stimuli is presented at least 1000 times. Therefore, it is likely that after several of hundred presentations of the same sounds, participants learn to distinguish the acoustic properties of speech sounds even if they originally belonged to the same phonetic category.

In the present investigation, five precautionary steps were followed to avoid this problem. First, a sequence of five standard and five deviant sounds was used. The goal was to reduce as much as possible a fixed acoustic pattern between standard and deviant stimuli. Second, the same standard sound did not occur two times consecutively and, there were at least three standard sounds with different VOT between the offset and the onset of a standard sound with equal VOT value. Therefore, the memory trace of the standard sounds was more likely to be a phonetic rather than an acoustic trace. Third, pairwise combinations between standards and deviants were fixed. That is, each deviant sound was preceded six times by each standard sound. The goal was to balance the acoustic differences between standards and deviants. Fourth, participants' ERPs were collected in 10 blocks of 75 s/each. It was expected that multiple short recordings, rather than one large recording would help to avoid participants learning the acoustic differences between speech sounds with equal phonetic membership. Additionally, participants had breaks of 1 to 1.5 min between blocks. The breaks were used to reduce participants' fatigue and to make them hear natural speech in accordance with the

language context. Finally, participants' ERPs were recorded in a single session rather than in two experimental sessions. This was done in order to control as much as possible for participants' state throughout the experimental session (Hillyard et al., 1971).

### **Do bilinguals and monolinguals show a MMN?**

The MMN typically is manifested in the deviant ERP as an amplitude increase in P2 and/or N2 components (Naatanen, 1992). Therefore, the evaluation of deviant P2 and N2 was done by comparing their peak amplitudes with baseline (Hillyard et al., 1973). It was expected that both monolinguals and bilinguals would show larger N2 and/or P2 amplitudes in the deviant response with respect to the standard response.

The results showed that both groups presented a change in amplitude in P2 and N2 components with respect to the standard, but P2 had a less positive response and N2 had a more negative response when compared with the standard ERP. The less positive values of P2 deviant component can be explained as the result of the overlap in time with the negative response of the MMN. In the same way, N2 component showed more negative values than its standard counterpart. The subtraction of the standard response from the deviant response showed that the time range when the MMN component was present was approximately from 150-350ms after stimulus onset (Naatanen, 1982). This confirms that the less positive and more negative values obtained in P2 and N2 deviant components, respectively, were the result of the MMN. Therefore, it can be concluded that the method implemented in this investigation was successful in generating the MMN.

### **Do bilinguals show different MMN amplitudes across language contexts?**

In the present investigation it was hypothesized that bilinguals, but not monolinguals, would show a change in the acoustic-to-phonemic representations among



language contexts. Bilinguals' phonemic boundary was expected to shift toward the voiced endpoint in the Spanish language context, and toward the voiceless endpoint in the English language context. The shift toward the voiced endpoint would result in the perception of more /t/ sounds than /d/ sounds. Therefore, the perception of more deviant sounds in the Spanish language context was hypothesized to produce an increase in pre-attentive discrimination (i.e., larger MMN amplitude). The results showed a significant amplitude change in the N2 deviant component between language contexts for bilinguals, but not for monolinguals. However, the change in amplitude across language context for N2 component was in the opposite direction of what was hypothesized. That is, bilinguals showed a more negative N2 peak in the English language context and less negative N2 amplitude in the Spanish language context. According to Tiitinen et al. (1994), the amplitude of the MMN increases as the ability to detect acoustic differences between standard and deviant increases. This suggests that bilinguals were more sensitive to phonetic differences in the English language context than in the Spanish language context.

The amplitude of the MMN showed the same amplitude trend that the N2 deviant component (larger amplitude in the English language set) in bilinguals, but not in monolinguals. However, in this case the amplitude of the MMN among language contexts did not reach significance. Two facts should be considered: 1) the MMN showed the same amplitude trend between language contexts in bilinguals with respect to N2 deviant ERP; and 2) bilinguals' MMN amplitude change among language contexts MMN showed a p-value of .11, whereas monolinguals showed a p-value of .99. These facts suggest that bilinguals' lack of statistical significance in the MMN between

language contexts was the result of individual variability in the difference between standard and deviant potentials. This issue may be resolved by increasing the sample size, but still more research is needed in order to clarify this outcome.

### ***Why ERP Amplitudes Went in the Opposite Direction of What was Expected***

It was expected that bilinguals in the Spanish language context would perceive more sounds as deviants than in the English language context. In other words, fewer number of sounds would be perceived as standards in the Spanish language than in the English language context. Thus, it was expected that larger N2 and MMN amplitudes would result in the Spanish language context due to larger number of deviants. But, as mentioned previously, this outcome was not observed. A closer examination indicates the results might not be as puzzling as they initially seem. The basis of MMN elicitation is the formation of a memory trace that is created from the standard sound being presented (Naatanen, 1992). It has been shown that the lower the probability a deviant will occur, the higher the amplitude of the MMN (Naatanen, Sams, Jarvilehto, & Soininen, 1983). Therefore, it is possible that larger N2 amplitude in the English language context was due to a perceptual decrease in the deviant proportion. In other words, in the English language context, bilinguals perceived fewer number of sounds as deviant yielding to a better memory trace for 'd' due to more deviants perceived as standards. Consequently, bilinguals formed a better 'd' memory trace in the English language context than in the Spanish language, and therefore, a stronger acoustic contrast between standard and deviant occurred. In contrast, in the Spanish language context, a lesser number of standard sounds functioned as 'd' memory traces. Therefore, acoustic contrasts were less salient and smaller amplitudes resulted. Moreover, it seems that the

more standards there were preceding a deviant sound, the larger N2/MMN amplitudes. These results suggest that perception of sounds was influenced by the language context; however, it does not confirm in a direct way the double phonemic boundary in bilinguals.

#### **CORRELATION BETWEEN BEHAVIOR AND ELECTROPHYSIOLOGY**

The third goal of the present investigation was to examine the relationship between behavioral and electrophysiological responses. An existing literature supports the notion that the MMN can be used to test phonetic categories (Nenonen et al., 2005; Peltola et al., 2003; Rivera-Gaxiola et al., 2000; Rivera-Gaxiola, et al., 2000b; Sharma & Dorman, 1998, 2000; Winkler et al., 1999). However, there is little or no research that has correlated shifts in the voicing boundary with electrophysiological findings. Therefore, an aim of the present investigation was to test whether bilinguals' double phonetic category correlated with the amplitude of the ERP components. It was expected that the degree in which the phonemic boundary shifted would correlate with the degree of the amplitude of the ERP components. For example, larger shifts in the phonemic boundary would be associated with larger amplitudes in the ERP components. Also, smaller shifts in the voicing boundary would be associated with smaller amplitudes in the ERP components.

The bilingual group, but not the monolingual group, showed a significant correlation. The correlation suggests that bilinguals' shifts in the voicing boundary correlated with the amplitude of the  $P2_p$ - $N2_v$  component. Bilinguals that showed a voicing boundary shift toward the voiceless endpoint in the English language context (expected direction) showed bigger  $P2_p$ - $N2_v$  amplitude. In contrast, bilinguals that showed a voicing boundary shifts in the opposite direction than the expected showed

smaller  $P2_p$ - $N2_v$  amplitude. Contrary to bilinguals, monolinguals showed no significant correlation between the voicing boundary and  $P2_p$ - $N2_v$  amplitude.

These findings support the claim that changes in amplitude are associated with the difference in proportions between standards and deviants presented (Naatanen et al., 1983). In the present investigation, the proportion of standard and deviant sounds was kept constant. However, the proportion of deviants sounds perceived differed among language conditions. For example, the significant positive correlation (see Figure 11) shows that those bilinguals who perceived fewer sounds as deviants had larger ERP amplitudes, whereas those bilinguals who perceived larger number of deviants had smaller ERP amplitudes. Interestingly, bilinguals, but not monolinguals show this amplitude effect. This finding confirms that only bilinguals' speech perception was influenced by the language context.

#### **LIMITATIONS AND FUTURE DIRECTIONS**

The variability encountered in bilingual research is a common problem (Elman et al., 1977; Flege & Eefting, 1987a; Hazan & Boulakia, 1993). Finding bilinguals who are equally fluent in both languages is the exception rather than the rule (Grosjean, 1982). Although it was paramount in this study to recruit a homogenous sample of fluent bilinguals in English and Spanish, there was some variability in the sample. For example, because bilinguals from this study live in the U.S. and are surrounded by an English context most of the time, this influenced the percentage of time they used English and Spanish. Indeed, when bilinguals were asked to indicate the percentage of time they use English and Spanish in their daily lives on a scale from 1 (i.e., 100% of the time Spanish) to 11 (i.e., 100% of the time English, see appendix A), the bilingual mean was 8.06 (*SD*

= 2.9). This suggests that bilinguals who participated in this investigation used more English than Spanish in general. This fact might have influenced the perception of speech sounds across language contexts.

In future studies it would be interesting to recruit Spanish-English bilinguals living in a Spanish-speaking country and observe how they perceive speech sounds across language contexts. In order to accomplish such a goal, one would need to recruit bilinguals who grew up in a Spanish-speaking country, but had parents whose native language is English, and thus learned English from their parents. Also, these bilinguals must use English in their daily lives, so perhaps participants studying in a bilingual school could be recruited. Finally, it would be necessary to recruit monolinguals of Spanish as a control group. Clearly, such a study would be helpful in understanding the bilingual brain.

One of the limitations of this study was that electrophysiological and behavioral responses were concurrently obtained during the perceptual task. Although this approach was implemented to obtain reliable ERP measurements, behavioral measurements could have been biased. For example, in order to increase the signal to noise ratio in brain responses obtained from the scalp, it is necessary to increase the number of stimulus presentations. But, as was discussed previously, this approach might have produced a range effect (Brady & Darwin, 1978; Keating et al., 1981). Thus, in this study while trying to implement an ideal electrophysiological methodology, the behavioral responses might have been compromised. It is important to note, however, that because ERPs and behavioral responses were measured concurrently, it was possible to correlate bilinguals' shift in the double boundary with electrophysiological responses.

It has been reported that some electrophysiological responses vary in accordance with participants' gender (McFadden & Champlin, 2000). The evidence in this field has shown that the amplitudes of AEP components in response to transient stimulation differ in accordance to sex preference and sex. Unfortunately, even though both males and females are recruited in speech categorization experiments, there are no studies that have analyzed sex differences. This might be due to the fact that most studies' sample size is too small to analyze differences across gender. As in other studies, this study's sample size was too small to make analysis by gender possible. It is clear that in future studies it would be informative to observe how speech perception varies, not only according to language usage, but also according to gender.

## **CONCLUSION**

The primary goal of the present investigation was to assess the double phonemic boundary in bilinguals. Behavioral and electrophysiological measures were obtained from bilingual speakers of Spanish and English and monolinguals speakers of English. It was found that neither bilinguals nor monolinguals showed a shift in the voicing boundary in accordance with the language set. It was speculated that the implementation of the odd-ball paradigm to collect behavioral responses may have caused a range-effect that obscured the double phonemic boundary in bilinguals. The analysis of speech tokens close to or at the phonemic boundary (ambiguous zone) showed that bilinguals, but not monolinguals, perceived less frequently deviant sounds in the English language context than in the Spanish language context. These results suggested that bilinguals' perception of speech sounds was influenced by the language context.

The ERP results varied according to the proportion of deviant sounds perceived among language contexts. Specifically, bilinguals, but not monolinguals, showed an increment in the N2/MMN ERP components' amplitude when less number of sounds were perceived as deviants (i.e., English language context). In the contrast, a reduction of the ERP components' amplitude was found when sounds were perceived more times as deviants (i.e., Spanish language context). The correlations between behavior and electrophysiology revealed that the increase in ERP amplitude was related to the number times bilinguals' pressed the response button across language context. These results suggest that while bilinguals' pre-attentive discrimination was influenced by the language context, it does not confirm in a direct way the double phonemic boundary in bilinguals.

This study was innovative because it attempted to examine how language use in bilinguals might influence the perception of sounds by using two different methodologies. Clearly, each of these methodologies had its limitations and more research needs to be done in this area. Nonetheless, the results support rather than contradict the existence of a double phonemic boundary in bilinguals.

## Appendix A: Language Questionnaire

Participant ID \_\_\_\_\_ Name: \_\_\_\_\_ Age \_\_\_\_\_

Field of Study \_\_\_\_\_

Year of Study \_\_\_\_\_

Date     /     /

Please answer and send the completed document via e-mail. Please type your personal information in the *above* and *below* shaded fields. After answering the questions, **SAVE** the document and send it to [gasa@austin.utexas.edu](mailto:gasa@austin.utexas.edu). Please send the document as an attachment only.

Where were you born? \_\_\_\_\_

How long have you lived in the US? \_\_\_\_\_

This questionnaire is related to the amount of Spanish and English you have been exposed in your life. Please select the box that best describe the percentage of Spanish or English you have been exposed in the given age. If you were exposed only to one language in a specific age, please select the 100% box for that language.



## SECTION 1: EXPOSURE

### QUESTION 1.

From the following years of age please select which language you heard, spoke and read the most. For example, if you indicate you heard English 75% of the times between 6-9 years of age, it means that you heard Spanish the remaining 25% of the times.

	L A N G U A G E Y O U <b>H E A R D</b> T H E M O S T				
	Spanish 100%	Spanish 75% English 25%	Spanish 50% English 50%	Spanish 25% English 75%	English 100%
<b>Age</b>					
0-3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3-6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6-9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9-12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12-15	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15-18	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18-21	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21-24	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24-27	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27-30	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30-33	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	L A N G U A G E Y O U <b>S P O K E</b> T H E M O S T				
	Spanish 100%	Spanish 75% English 25%	Spanish 50% English 50%	Spanish 25% English 75%	English 100%
<b>Age</b>					
3-6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6-9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9-12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12-15	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15-18	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18-21	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21-24	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24-27	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27-30	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30-33	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	L A N G U A G E   Y O U <b>R E A D</b> T H E   M O S T				
	Spanish 100%	Spanish 75% English 25%	Spanish 50% English 50%	Spanish 25% English 75%	English 100%
Age					
9-12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12-15	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15-18	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18-21	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21-24	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24-27	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27-30	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30-33	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## Appendix A: Continue

### SECTION 2: CONFIDENCE

In reference to ENGLISH, select the option that best describes you. Please select only one box per question.

#### Question 1      IN    reference to    **E N G L I S H**

- ☐ I cannot speak the language, I have a few words or phrases and, I cannot produce sentences
- ☐ I cannot speak the language, I have some words or phrases and,  
**I can produce some sentences**
- ☐ I have a limited proficiency with grammatical errors and a limited vocabulary
- ☐ I have good proficiency with some grammatical errors and I have some social and academic vocabulary
- ☐ I have a native-like proficiency with few grammatical errors and I have a good vocabulary

#### Question 2      IN    reference to    **E N G L I S H**

- ☐ I only understand a few words of what is being said
- ☐ I understand a vague idea of what is being said
- ☐ I understand the general idea of what is being said
- ☐ I understand most of what is being said
- ☐ I understand all of what is being said

In reference to SPANISH, select the option that best describes you. Please select only one box per question.

**Question 3**      IN    reference to    **S P A N I S H**

- ☐ I cannot speak the language, I have a few words or phrases and, I cannot produce sentences
- ☐ I cannot speak the language, I have some words or phrases and, **I can produce some sentences**
- ☐ I have a limited proficiency with grammatical errors and a limited vocabulary
- ☐ I have good proficiency with some grammatical errors and I have some social and academic vocabulary
- ☐ I have a native-like proficiency with few grammatical errors and I have a good vocabulary

**Question 4**      IN    reference to    **S P A N I S H**

- ☐ I only understand a few words of what is being said
- ☐ I understand a vague idea of what is being said
- ☐ I understand the general idea of what is being said
- ☐ I understand most of what is being said
- ☐ I understand all of what is being said

From the following age ranges please indicate with which language you were most confident when speaking, hearing and reading it. Confidence does not mean the language you used the most. For example, it might be possible that between 9-12 years of age you heard English at school and Spanish at home. However, you felt more self-confident when hearing Spanish than when hearing English. If you were exposed to only **one language** in a specific age, answer for the **exposed** language only.

		CONFIDENCE IN HEARING				
		Not confident	25% confident	50% confident	75% confident	Strong confident
Age	Language					
9-12	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12-15	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15-18	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18-21	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21-24	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24-27	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27-30	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30-33	English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		CONFIDENCE IN SPEAKING							
		Not confident	25% confident			50% confident	75% confident		Strong confident
Age	Language								
9-12	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
12-15	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
15-18	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
18-21	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
21-24	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
24-27	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
27-30	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
30-33	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

		CONFIDENCE IN READING							
		Not confident	25% confident			50% confident	75% confident		Strong confident
Age	Language								
9-12	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
12-15	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
15-18	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
18-21	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
21-24	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
24-27	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
27-30	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
30-33	English	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
	Spanish	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

### SECTION 3: LANGUAGE USAGE IN EVERYDAY LIVES

3. Choose the percentages that best describe how fluent your parents and siblings are in English and Spanish (pop-down menu)

	English	Spanish	It doesn't apply
<b>Mother</b>	0%	0%	<input type="checkbox"/>
<b>Father</b>	0%	0%	<input type="checkbox"/>
<b>Siblings</b>	0%	0%	<input type="checkbox"/>

4. Type the years your parents have lived in United States and/or in a Spanish speaking country.

If they have lived all their life in one country please choose US or Spanish Speaking country (pop-down menu).

	Father		Mother
United States	<input type="text"/>	United States	<input type="text"/>
Select Country	<input type="text"/>	Select Country	<input type="text"/>
All their life in	---	All their life in	---
It doesn't apply	<input type="checkbox"/>	It doesn't apply	<input type="checkbox"/>

5. Estimate in which percentage you use Spanish and English with (please mark the percentage that fits better, if it doesn't apply please indicate it with an X).

		Spanish					50% Spanish 50% English					English
	It doesn't apply	100%	90%	80%	70%	60%	50%	60%	70%	80%	90%	100%
1. Employers/teachers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Mother	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Father	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Siblings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Partner	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Yourself	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Classmates/peers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Pets/plants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## Appendix B: Stimulus Sequence

Recording Blocks									
1	2	3	4	5	6	7	8	9	10
1	2	76	2	151	1	226	1	301	4
2	1	77	4	152	2	227	4	302	5
3	4	78	1	153	4	228	5	303	1
4	5	79	5	154	3	229	2	304	2
5	7	80	2	155	5	230	3	305	4
6	2	81	8	156	10	231	9	306	7
7	1	82	1	157	2	232	1	307	3
8	5	83	4	158	3	233	4	308	1
9	10	84	5	159	1	234	2	309	5
10	4	85	3	160	8	235	6	310	4
11	1	86	2	161	2	236	3	311	10
12	3	87	10	162	4	237	5	312	1
13	5	88	1	163	5	238	2	313	5
14	2	89	5	164	8	239	1	314	4
15	7	90	3	165	3	240	4	315	2
16	4	91	6	166	4	241	5	316	8
17	5	92	2	167	1	242	8	317	5
18	2	93	4	168	2	243	3	318	4
19	8	94	3	169	5	244	2	319	3
20	4	95	7	170	6	245	5	320	2
21	5	96	2	171	1	246	1	321	1
22	1	97	1	172	2	247	7	322	9
23	3	98	3	173	3	248	3	323	3
24	9	99	4	174	5	249	4	324	4
25	2	100	5	175	1	250	5	325	1
26	4	101	9	176	4	251	9	326	2
27	3	102	2	177	7	252	2	327	8
28	10	103	3	178	3	253	3	328	5
29	5	104	4	179	2	254	4	329	3
30	1	105	7	180	4	255	5	330	2
31	3	106	1	181	1	256	8	331	6
32	2	107	5	182	9	257	2	332	1
33	4	108	3	183	2	258	1	333	4
34	6	109	9	184	3	259	5	334	3
35	5	110	4	185	5	260	4	335	5
36	1	111	1	186	1	261	10	336	9
37	3	112	3	187	6	262	1	337	1
38	8	113	8	188	4	263	5	338	3
39	2	114	5	189	5	264	4	339	4
40	4	115	4	190	2	265	7	340	2
41	1	116	1	191	9	266	1	341	7
42	10	117	3	192	4	267	5	342	1
43	5	118	6	193	1	268	2	343	4
44	3	119	4	194	5	269	6	344	2
45	4	120	2	195	9	270	3	345	3
46	2	121	1	196	3	271	5	346	5
47	8	122	10	197	4	272	4	347	1
48	3	123	3	198	2	273	8	348	10
49	4	124	5	199	1	274	1	349	2
50	5	125	2	200	7	275	3	350	3
51	7	126	4	201	4	276	2	351	4
52	1	127	8	202	3	277	7	352	9
53	4	128	1	203	5	278	1	353	5
54	2	129	5	204	1	279	3	354	1
55	6	130	2	205	10	280	2	355	3
56	5	131	9	206	2	281	9	356	4
57	3	132	1	207	5	282	5	357	6
58	2	133	3	208	4	283	4	358	5
59	1	134	4	209	1	284	3	359	2
60	4	135	6	210	3	285	10	360	3
61	9	136	2	211	6	286	5	361	1
62	3	137	5	212	4	287	2	362	8
63	5	138	1	213	2	288	1	363	5
64	2	139	4	214	5	289	4	364	4
65	4	140	10	215	3	290	3	365	2
66	1	141	5	216	4	291	6	366	3
67	3	142	3	217	8	292	1	367	7
68	9	143	4	218	5	293	4	368	5
69	5	144	1	219	1	294	2	369	2
70	1	145	5	220	3	295	3	370	1
71	3	146	7	221	10	296	1	371	6
72	6	147	2	222	5	297	4	372	3
73	2	148	3	223	2	298	10	373	2
74	1	149	5	224	3	299	3	374	5
75	3	150	2	225	7	300	2	375	10

VOT  
code

Standards				
-20	-15	-10	-5	0
1	2	3	4	5

VOT  
code

Deviants				
5	10	15	20	25
6	7	8	9	10



## References

- Abramson, A. S., & Lisker, L. (1967). *Discriminability along the voicing continuum: Cross language tests*. Paper presented at the Proc. 6th Int. Congr. of Phonet. Sci., Prague.
- Abramson, A. S., & Lisker, L. (1972). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, 1, 1-8.
- Altenmuller, E. O., & Gerloff, C. (1998). Psychophysiology and the EEG. In E. Niedermeyer, & F. Lopes Da Silva (Eds.), *Electroencephalography: Basic principles, clinical applications, and related fields* (pp, 637-655). New York: Lippincott Williams & Wilkins.
- Balota, D. A., & Duchek, J. M. (1986). Voice-specific information and 20-seconds delayed-suffix effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 12, 509-516.
- Benet-Martínez, V., & John, O. P. (1998). Los Cinco Grades across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, 75, 729-750.
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345-360.
- Bohn, O. S., & Flege, J. E. (1993). Perceptual Switching in Spanish-English Bilinguals. *Journal of Phonetics*, 21(3), 267-290.
- Burlingame, E., Sussman, H. M., Gillam, R. B., & Hay, J. F. (2005). An investigation of speech perception in children with specific language impairment on a continuum of formant transition duration. *Journal of Speech Language and Hearing Research*, 48(4), 805-816.
- Brady, S. A., & Darwin, C. J. (1978). A range effect in the perception of voicing. *Journal of the Acoustical Society of America*, 63, 1556-1558.
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustic Society of America*, 54(2).
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., et al. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, 1, 351-353.
- Cowan, N. (1995). *Attention and Memory. An integrated framework*. Oxford, England: Oxford University Press.

- Diamond, D. M., & Weinberger, N. M. (1986). Classical-Conditioning Rapidly Induces Specific Changes in Frequency Receptive-Fields of Single Neurons in Secondary and Ventral Ectosylvian Auditory Cortical Fields. *Brain Research*, 372(2), 357-360.
- Diehl, R. L., Elman, J. L., & McCusker, S. B. (1978). Contrast effects on stop consonant identification. *Human Perception and Performance*, 4(4), 599-609.
- Donchin, E., Ritter, W., & McCallum, W. C. (1978). Cognitive Psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, and S. H. Koslow, (Eds.), *Event Related Potentials in the man*. New York: Academy Press
- Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 4, 99-109.
- Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of Acoustical Society of America*, 62(4), 971-974.
- Flege, J. E., & Eefting, W. (1987a). Cross-Language Switching in Stop Consonant Perception and Production by Dutch Speakers of English. *Speech Communication*, 6(3), 185-202.
- Flege, J. E., & Eefting, W. (1987b). Production and Perception of English Stops by Native Spanish Speakers. *Journal of Phonetics*, 15(1), 67-83.
- Gao, E., & Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proceedings of the National Academy of Sciences of the United States of America*, 95(21), 12663-12670.
- Garcia-Sierra, A., & Champlin, C. A. (2003). Shift in the perceived voicing boundary of bilingual listeners. *Journal of the Acoustical Society of America*, 114(4), 2336.
- Garcia-Sierra, A., & Champlin, C. A. (2004). Bilinguals' categorical perceptual shift when producing Spanish or English words: Electrophysiological correlates. *Journal of the Acoustical Society of America*, 115(5), 2466.
- Greenberg, S., Marsh, J. T., Brown, W. S., & Smith, J. C. (1987). Neural Temporal Coding of Low Pitch .1. Human Frequency-Following Responses to Complex Tones. *Hearing Research*, 25(2-3), 91-114.
- Grosjean, F. (1982). *Life with two languages*. Cambridge, MA: Harvard University Press.
- Hall, J. W. III. (1992). *Handbook of Auditory Responses*. Allyn and Bacon. Boston.
- Hazan, V. L., & Boulakia, G. (1993). Perception and Production of a Voicing Contrast by French-English Bilinguals. *Language and Speech*, 36, 17-38.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, Vol. 182(4108), 177-179.

- Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science*, Vol. 172(3990), 1357-1360.
- Holt, L. L. (2005). Temporally nonadjacent nonlinguistic sounds affect speech categorization. *Psychological Science*, 16(4), 305-312.
- Holt, L. L., & Lotto, A. J. (2002). Behavioral examinations of the level of auditory processing of speech context effects. *Hearing Research*, 167(1-2), 156-169.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2<sup>nd</sup> Ed.). (pp. 102-138). New York: Guilford Press.
- Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, 60, 286-319.
- Keating, P. A., Mikos, M. J., & Ganong, W. F. (1981). A Cross-Language Study of Range of Voice Onset Time in the Perception of Initial Stop Voicing. *Journal of the Acoustical Society of America*, 70(5), 1261-1271.
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67(3), 971-990.
- Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50(2), 93-107.
- Kuhl, P. K. (1993). Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics*, 21(1-2), 125-139.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by six- months of age. *Science*, 255, 606-608.
- Liberman, A. M., Harris, K. S., Kinney, J. A., & Lane, H. (1961). The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology*, 61, 379-388.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Lotto, A. J., Sullivan, S. C., & Holt, L. L. (2003). Central locus for nonspeech context effects on phonetic identification. *Journal of the Acoustical Society of America*, 113, 53-56.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (Second ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

- McFadden, D., & Champlin, C. (2000). Comparison of Auditory Evoked Potentials in Heterosexual, Homosexual, and Bisexual Males and Females. *Journal of the Association for Research Otolaryngology*, 1(1), 89-99.
- Naatanen, R. (1982). Processing Negativity - an Evoked-Potential Reflection of Selective Attention. *Psychological Bulletin*, 92(3), 605-640.
- Naatanen, R. (1986). A classification of N2 kinds of ERP-components. In Cerebral psychophysiology: Studies in event related potentials, In W.C. McCallum, R. Zappoli, & F. Denoth (Eds.), *Electroencephalography & Clinical Neurophysiology, Suppl 38* (169-172).
- Naatanen, R. (1992). *Attention and Brain Function*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Naatanen, R., Gaillard, A. W. K., & Mantysalo, S. (1978). Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42, 313-329.
- Naatanen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, 42(1), 25-32.
- Naatanen, R., Lehtokoski, A., Lennes, M., & et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385, 432-434.
- Naatanen, R., & Michie, P. T. (1979). Early Selective-Attention Effects on the Evoked-Potential - a Critical-Review and Reinterpretation. *Biological Psychology*, 8(2), 81-136.
- Naatanen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*, 115(1), 140-144.
- Naatanen, R., Sams, M., Jarvilehto, T., & Soininen, K. (1983). Probability of deviant stimulus and event-related potentials. In R. Sinz, & M. R. Rosenzweig (Eds.), *Psychophysiology 1980* (pp. 397-405). Jean: VEB Gustav Fischer Verlag, and Amsterdam: Elsevier Biomedical Press.
- Naatanen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology*, 14(1-2), 53-98.
- Naatanen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125(6), 826-859.
- Nenonen, S., Shestakova, A., Huotilainen, M., & Naatanen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain and Language*, 92(1), 26-32.

- Osterhout, L., McLaughlin, J., & Bersick, M. (1997). Event-related brain potentials and human language. *Trends in Cognitive Sciences*, 1(6), 203-209.
- Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Naatanen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neuroscience Letters*, 352(1), 25-28.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., et al. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience*, 12(6), 1038-1055.
- Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials: II. Effects of attention. *Electroencephalography & Clinical Neurophysiology*, Vol. 36(2), 191-200.
- Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials: I. Evaluation of components. *Electroencephalography & Clinical Neurophysiology*, Vol. 36(2), 179-190.
- Rivera-Gaxiola, M., Csibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000a). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioral Brain Research*, 111(1), 13-23.
- Rivera-Gaxiola, M., Johnson, M. H., Csibra, G., & Karmiloff-Smith, A. (2000b). Electrophysiological correlates of category goodness. *Behavioral Brain Research*, 112(1-2), 1-11.
- Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., & Kuhl, P. K. (2005). Neural patterns to speech and vocabulary growth in American infants. *Neuroreport: For Rapid Communication of Neuroscience Research*, 16(5), 495-498.
- Scharf, B., & Houstsna, A. J. (1986). *Audition: II. Loudness, pitch, localization, aural distortion, pathology* (Vol. 1. Sensory processes and perception). New York: Wiley.
- Sharma, A., & Dorman, M. F. (1998). Exploration of the perceptual magnet effect using the mismatch negativity auditory evoked potential. *Journal of the Acoustical Society of America*, 104, 511-517.
- Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *Journal of the Acoustical Society of America*, 107(5), 2697-2703.
- Suga, N., Gao, E. Q., Zhang, Y. F., Ma, X. F., & Olsen, J. F. (2000). The corticofugal system for hearing: Recent progress. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), 11807-11814.
- Sussman, E., Winkler, I., Huotilainen, M., Ritter, W., & Naatanen, R. (2002). Top-down effects can modify the initially stimulus-driven auditory organization. *Cognitive Brain Research*, 13(3), 393-405.

- Tiitinen, H., May, P., Reinikainen, K., & Naatanen, R. (1994). Attentive Novelty Detection in Humans Is Governed by Pre-Attentive Sensory Memory. *Nature*, 372(6501), 90-92.
- Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception and Psychophysics*, 21(4), 289-297.
- Winkler, I., Kujala, T., Alku, P., & Naatanen, R. (2003). Language context and phonetic change detection. *Cognitive Brain Research*, 17(3), 833-844.

## **Vita**

Adrián García Sierra was born October 31st, 1973 in Mexico City, Mexico to Carlos García López and Gisela Sierra Otero and younger brother to Carlos García Sierra. Adrián grew up in Mexico City and attended the Institute of Fine Arts and Sciences. He then completed his B.A. in Psychology, in the University of the Americas in Puebla, Mexico. From 1999-2002 he conducted research with Maritza Rivera Gaxiola, at the Center of Neurobiology in Queretaro Mexico, and then at the Institute for Learning and Brain Sciences at the University of Washington along with Patricia Kuhl. In the fall of 2002, he began his Ph.D. in Speech and Hearing Sciences at The University of Texas at Austin under the guidance of Craig Champlin. During graduate school he conducted research with Dennis McFadden, Randy Diehl, and Harvey Sussman. In 2005 he received the US-Mexico/Borderlands Research Award, and in 2006 he was awarded with the Jesse H. Jones Fellowship and the Jennifer Sparrgrove Endowment Fellowship. Upon graduation, he will be assuming a Research Associate position at the University of Washington, Seattle.

Permanent address: 3456 North Hills Dr. 207c, Austin, TX, 78731

This dissertation was typed by the author.